

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2005, 03:37:22 ; Search time 4424 Seconds
(without alignments)
9963.477 Million cell updates/sec

Title: US-09-998-059-1

Perfect score: 1158
Sequence: 1 cacaacatacactcaaat.....cttcacaacaacaaag 1158

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsat:*
9: gb_gsat2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	395.4	34.1	413	8	BH911706
2	333.4	28.8	400	8	BZ289690
3	145.2	12.5	692	8	BZ084475
4	129.4	11.2	650	8	BH655922
5	83	7.2	963	8	CL515027
6	64	5.5	987	9	CNS014PQ
7	61.6	5.3	1101	9	CNS017V2
8	60.6	5.2	928	9	CNS00DKY
9	60.4	5.2	1101	9	CNS016L1
10	60.2	5.2	922	9	CNS0073W
11	60	5.1	759	9	CNS06QXV
12	59.4	5.1	324	9	CNS0243V
13	59.4	5.1	1187	9	CL648142
14	59.2	5.1	1092	9	CNS020K7
15	58.8	5.1	1045	9	CNS03YR4
16	58.4	5.0	1001	9	CNS01400
17	57.8	5.0	1101	9	AL107814
18	57.6	5.0	627	1	AL1525005
19	57	4.9	873	9	CNS024M5
20	57	4.9	987	9	CNS014PQ
21	57	4.9	1039	9	CNS006H7
22	56.8	4.9	1146	9	CNS021G2
23	56.8	4.9	555	5	BQ739452
24	56.8	4.9	583	9	CNS044QG

25	56.6	4.9	924	9	CNS07A5L	AL436159	T7 end of
26	56.6	4.9	982	9	CNS004WF	AL055517	Drosophila
27	56.4	4.9	764	7	CF289381	CF289381	AGENCOURT
28	56.4	4.9	886	8	BH177277	BH177277	008 L 22-
29	56.4	4.9	886	9	CNS07JUX	AL614235	T3 end of
30	56.4	4.9	966	9	CNS006H7	AL055882	Drosophila
31	56.2	4.9	912	8	AZ531076	AZ531076	ENTB023TR
32	56.2	4.9	911	8	AZ691769	AL065759	ENTH197F
33	56	4.8	1001	9	CNS006MN	AL050523	Drosophila
34	56	4.8	1001	9	CNS0155H	AL105023	Drosophila
35	55.8	4.8	607	7	CV490737	CV490737	AGENCOURT
36	55.8	4.8	892	9	CL101621	CL101621	ISB1-38N5
37	55.8	4.8	1101	9	CNS001FB	AL060732	Drosophila
38	55.8	4.8	2300	3	CR734152	CR734152	Tetradon
39	55.6	4.8	920	9	CNS0062R	AL061710	Drosophila
40	55.4	4.8	843	9	CNS00CS1	AL059666	Drosophila
41	55.2	4.8	839	9	AG518301	AG518301	Mus muscu
42	55.2	4.8	928	9	CNS00DKY	AL071865	Drosophila
43	55.2	4.8	1101	9	CNS001L72	AL078714	Drosophila
44	55.2	4.8	1201	9	CNS0168L	AL106431	Drosophila
45	55	4.7	888	9	CNS007VQ	AL050836	Drosophila

ALIGNMENTS

RESULT 1
LOCUS BH911706/c
DEFINITION SALK_071668.55.25.x Arabidopsis thaliana T7 end of
survey sequence.

ACCESSION BH911706.1 GI:22724639
VERSION BH911706
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 413)
Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R.,
Gadriab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L.,
Shim, P., Zimmerman, J., and Ecker, J.R.
A sequence-indexed library of insertion mutations in the
Arabidopsis Genome
Unpublished (2001)

JOURNAL CONTACT: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu

COMMENT This is single pass sequence recovered from the left border of
TDNA.
Class: TDNA tagged.

FEATURES Location/Qualifiers
1..413

source /organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Col-0"
/db_xref="taxon:3702"
/clone="SALK_071668.55.25.x"
/clone_lib="Arabidopsis thaliana T7 end of
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN

/note="vector: potw13; Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Bressani oleracea T01000DH3 buds provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Fabio Kahlnowicz (CSHL) and the shotgun library prepared at Washington University Genome Sequencing Center."

ORIGIN

	Query Match	12.5%	Score 145.2	DB 8	Length 692
	Best Local Similarity	69.8%	Pred. No. 2.9e-22		
	Matches	268	Conservative	0	Mismatches 108; Indels 8; Gaps 5
QY	770	ATACGTGAACACTGGTGAATCTTAATTTGTTGGTTAATGTT-TAAATTTAGATTCTATTCT	828		
Db	195	ATAAGGGGATCTAGTGAATGTAATTTGGTAACTTAATGTCAAAAAGTAGATTTCTCCAT	254		
QY	829	CCGGTTTAAAGTGAATATATATGTATCATGTGTTAAACATTTGTAAGTAAAGTAATTA	888		
Db	255	CCCCCTTAATTAATGAGTTAAATGAATTTAAG--ACGATTAATAAATGAATTAATCAACT	311		
QY	889	AATGATAAATTTAGTTGATGATGAATACTGGAACAAAAAATGAGATAGATTACATTTGATTT	948		
Db	312	AGAGTTCTCTGTACTGGTTGTATTAACATGAACCTTAAATATGAGATACAT-TTGAATTT	370		
QY	949	TGTCGTATTTTGACATATATCCGAGAGTAGTACTACGCCGATGAAATCAAGACACTTGC	1008		
Db	371	TGTCGTGTTTGGAGTATGCGAGATAGTGTCTACCCGATGAACATCATGAGACACTTGC	430		
QY	1009	TCGAGCTCACAGAGTACGTGTAAAAACCTTAGACTGAAGTCCCATGCAAACTCAATCC	1068		
Db	431	TTCAAGCTCACAGATGACGTGTAAAGACCATAGACCCACGACTTATGCAAAACAGTTCC	490		
QY	1069	TACGTGCTCAAAACGACAGCTCACTTGACATATATATAACTCTCTTAAGTCCGTTTC	1127		
Db	491	TACGTGGCAAAACCTTCATGTCTCACTCCAC-ATATATAAACTCCACCAAGTCTTCATCT	549		
QY	1128	-TCTTCAATCCATCTCTCAACAA 1150			
Db	550	TTCTTCATCCATCTATCAAAAAA 573			

RESULT 4
LOCUS BH655922
DEFINITION BH655922 650 bp DNA linear GSS 19-FEB-2002
BOHX116TR_BO_2_3_KB Brassica oleracea genomic clone BOHX16,
genomic survey sequence.
ACCESSION BH655922
VERSION BH655922.1 GI:18714215
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euarysta; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots;
rosoids; eurooids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 650)
AUTHORS Town,C.D., Van Aken,S., Uteerack,T., Koo,H. and Frazer,C.M.
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL Unpublished (2001)
COMMENT Contact: Chris Town

```

/db_xref="taxon:3712"
/clone="BOHx116"
/clone_lib="BO_2_3_KB"
/notes="Vector: pHOs1; site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOs1 using BstXI linkers

```

ORIGIN

Query Match	11.2%	Score 129.4	DB 8	Length 650
Best Local Similarity	77.9%	Pred. No. 1.1e-18		
Matches 194	Conservative	0	Mismatches 51	Indels 4
				Gaps 3
QY	904	TGATGATTAACGCTGAACCAAAAAATGAGATAGATACATTTGATTTTTCGTATTTTGACA	963	
Db	9	TGTTGATTAACATGTAATCTAAAAATGAGATACA-TTTGATTTTTCGTGTTGGAG	67	
QY	964	TATTCGGAGAGTGTAGCTTACGCGCATGAAGATCAAGAGACACTTGTCTGAGCTCAAGAGT	1023	
Db	68	TATCGGAGATATAGTCTACGCGCATGAACATCATGAGACACTTCTTCAGCTCAGAGT	127	
QY	1024	GACCTGTAAAAAGTTTGACTGAGTCCCATGCGAAACCTTAATCTCTACGTGGCTCAAAAC	1083	
Db	128	GACCTGTAAAGACCATGTAGACCCAGACTTATGCAAAACCAAGTCTCTACGTGGCAAAACCT	187	
QY	1084	ACGAGCTCACTTGACATATATATAATCCCTCTTAAGCCGTTCC--ACTTGATCATCATCTC	1141	
Db	188	TCATGCTCACTTCCAC-ATATATAAATCTCAACCAAGTCCATCTTTCTTCATCATCTCA	246	
QY	1142	TCACACACAA	1150	
Db	247	TCACAAAAA	255	

RESULT 5	CL515027/c	963 bp	DNA	linear	GSS 01-APR-2004
LOCUS	CL515027/c				
DEFINITION	SAIL_895_F04.v1 SAIL Collection Arabidopsis thaliana genomic clone SAIL_895_F04.v1, genomic survey sequence.				
ACCESSION	CL515027				
VERSION	CL515027.1	GI:46012347			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 963)				
AUTHORS	Sessions,A., Burke,E., Presting,G., Au,G., McElver,J., Patton,D., Dietrich,B., Ho,P., Baccand,J., Ko,C., Clarke,J.D., Cotton,D., Bullis,D., Snell,J., Miguel,T., Hutchinson,D., Kimerly,B., Mitzel,T., Katagiri,F., Glazebrook,K., Law,M. and Goff,S.A.				
TITLE	A high-throughput Arabidopsis reverse genetics system				
JOURNAL	Plant Cell 14 (12), 2985-2994 (2002)				
MEDLINE	22356987				
PubMed	12468722				
COMMENT	Contact: Sessions A				

3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number C854039; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: TDNA tagged.

FEATURES

```
source
1. .963
  /organism="Arabidopsis thaliana"
  /mol_type="genomic DNA"
  /ecotype="Columbia"
  /db_xref="taxon:3702"
  /clone="SAIL_895_F04.v1"
  /clone_11b="SAIL_Collection"
```

/note="T-DNA left border sequences were isolated using a modified TAIL-PCR strategy"

ORIGIN

Query Match 7.2%; Score 83; DB 9; Length 963;
Best Local Similarity 57.2%; Pred. No. 3.8e-08;
Matches 191; Conservative 0; Mismatches 137; Indels 6; Gaps 3;

97 GATTCAGACACAAAGCAGTAAGATAAGAAATTTAAAG- AACGCTATGTAGTGC 155
416 GTTTCAGGCTATCGGCTGTATGATGCAATGTGGCGCGCTCTGTGTAGTGTC 357
156 G-CAAAATATCTCTTAATCAAAACGTAACAAAGTAATTTGCAAAATCCGACAGAA 214
356 GCCGAAATGCTTTCTTGTCTTAGCAGTACCATCTGGATGACTATATCTTAGCAATTA 297
215 ACTCTCAGCCACCTCGGAATTCAGCTCTTCACTAAATTTTGAAGAATGATCAAT 274
236 AGCTTCGTCGCCGTCGAAATTTCACTTTTCAATTAATTTTGGAAAGTAATTTCTTT 237
275 ACCAACCATTACACAAATACATATCAAAATGCGCAATCG- ---TACCTGAAACT 330
236 ACTNTNTCATNTNNNNNNNNNNNTTGTAGATCACAGAAAGCAAGAACCGGAGCT 177
331 TTGCTTCAAGTGCAGAGAGAGAAAGAGATCTGAGAGAAAGGCTTTAGGTTTA 390
176 GCTTGTGGGTTCCCGCCCAATGCAAGTAATGTCAGCAACGCGGTAGGTTTAA 117
391 GCTCAGACTTATTTGAGTAATGAGACGGTGT 424
116 GCTCAGGCTTCCCTTGAGTAATGAGACGGGT 83

RESULT 6

CNS014PQ/c

987 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence Sp6 end of BAC

BACN12P22 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL104456
AL104456.1 GI:5616067

GSS.
Drosophila melanogaster (fruit fly)

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 987)

Genoscope.
Direct Submission

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a

collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC

library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector

pBelobAC11.
Location/Qualifiers

1..987
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"

/db_xref="taxon:7227"
/clone="BACN12P22"

/clone_1b="DrosBAC"
/plasmid="pBelobAC11"

/note="end : SP6"

Query Match 5.5%; Score 64; DB 9; Length 987;

Best Local Similarity 33.0%; Pred. No. 0.00076;
Matches 92; Conservative 69; Mismatches 118; Indels 0; Gaps 0;

681 TTTTATTTTATTTTATAGTCTTTGTTTATCTTCAATACAAATTTTGGCTGAT 740
954 WWWWWWTT 895
741 CTGCAACCTCTTCGATCATATCCCAATATAGCGAACTGTGATCTAATTTGTT 800
894 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 835
801 GTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 860
834 AAMNTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 775
861 TAAACATTTGATGATGATATATATATATATATATTTTATTTTATTTTATTT 920
774 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 715
921 CAAAAATGAGATGATATCATTTGATTTTGTGCTATTTT 959
714 AAAMTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 676

RESULT 7

CNS017V2

1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence Sp6 end of BAC

BACN37L10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.

AL108536
AL108536.1 GI:5628840

GSS.
Drosophila melanogaster (fruit fly)

Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)

Genoscope.
Direct Submission

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr

- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a

collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC

library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector

pBelobAC11.
Location/Qualifiers

1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"

/db_xref="taxon:7227"
/clone="BACN37L10"

/clone_1b="DrosBAC"
/plasmid="pBelobAC11"

/note="end : SP6"

Query Match 5.3%; Score 61.6; DB 9; Length 1101;
Best Local Similarity 44.8%; Pred. No. 0.0027;
Matches 129; Conservative 24; Mismatches 133; Indels 0; Gaps 0;

672 TTTCTCTGTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 731
761 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 820
732 TGGCTATCTTGCAAATCTTTCATATATCGCAATATACGTGAACACGTGATCTA 791

[illegible][illegible]

```

Db      1020 ADATTAMDAATRTAAATWTTTKTTTKTTTKTTTKTAKTKAKAGRAMWWW 1073

RESULT 10
CNS0073W/C      922 bp      DNA      linear      GSS 03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION      BACR14D09 of RPCL-98 library from Drosophila melanogaster (fruit
Fly), genomic survey sequence.
ACCESSION      AL066784.1 GI:4945247
VERSION      GSS.
KEYWORDS      Drosophila melanogaster (fruit fly)
SOURCE      Drosophila melanogaster
ORGANISM      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 922)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT      Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosagawa and
Aaron Mamoser in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCL-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridisation from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source      1..922
             /organism="Drosophila melanogaster"
             /mol_type="genomic DNA"
             /db_xref="taxon:7227"
             /clone="BACR14D09"
             /clone_11b="RPCL-98"
             /note="end : TET3"

ORIGIN
Query Match      5.2%; Score 60.2; DB 9; Length 922;
Best Local Similarity 25.5%; Pred. No. 0.0055;
Matches 96; Conservative 11; Mismatches 170; Indels 0; Gaps 0;

1 CACAAACAGTACCTCAAAATCCAGCTCACTCTCACTCAATATGCAACTTCATCATGAA 60
Db      819 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 760
Qy      61 AACATCAAAAACATCAAGTAACTCAAGTTCAGATTCAGACACAAAGCCAGTAAA 120
Db      759 CCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCC 700
Qy      121 GATAGAAAATTTAAGAACGCTCATGCTAGCTGGCGCAAAATCTTCTTAATCAAAACG 180
Db      699 MCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCC 640
Qy      181 TAACAACGAGTAATTAAGCAAAATCCGAGGAGAAAATCTCAACCCGCTCCGAATTCACG 240
Db      639 AAAAAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMA 580
Qy      241 TCTTCATAAATTTTCGAAAGAAATCGATCAATACCAACCATTTACAAATATCATATA 300
Db      579 MMCCASAMMAAMMAAMMAAACAAMMAACAASAAAAMMAAMMAAMMAAMMAAMMAAAA 520
Qy      301 TCAAAATGCGCAGAAATCGTACTGGAACCTTGCTTCAAGTCGACAGAGAGAGAAAAGA 360

```

```

Db      519 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCACAGRGCGGAGRGCGGGG 460
Qy      361 AGATCGTGAGAAAGG 377
Db      459 GGARRRAGGAGRAGRGR 443

RESULT 11
CNS06QXV/C      759 bp      DNA      linear      GSS 05-JUL-2001
LOCUS      T7 end of clone AM0A009H09 of library AM0A from strain CLIB 89 of
DEFINITION      Yarrowia lipolytica, genomic survey sequence.
ACCESSION      AL411257.1 GI:12180512
VERSION      GSS.
KEYWORDS      Yarrowia lipolytica
SOURCE      Yarrowia lipolytica
ORGANISM      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Dipodascaceae; Yarrowia.
REFERENCE      1 (bases 1 to 759)
AUTHORS      Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Boivin-Fukuhara,M., Bon,E., Brotier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowska-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE      Genomic exploration of the hemiascomycetous yeasts : 1. A set of
Yeast species for molecular evolution studies
JOURNAL      FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE      20584711
PUBMED      11152876
REFERENCE      2 (bases 1 to 759)
AUTHORS      Casaregola,S., Neuvéglise,C., Lepingle,A., Bon,E., Feynerol,C.,
Artiguenave,F., Wincker,P. and Gaillardin,C.
TITLE      Genomic exploration of the hemiascomycetous yeasts : 17. Yarrowia
lipolytica
JOURNAL      FEBS Lett. 487 (1), 95-100 (2000)
MEDLINE      20584727
PUBMED      11152892
REFERENCE      3 (bases 1 to 759)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT      This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
anusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
source      1..759
             /organism="Yarrowia lipolytica"
             /mol_type="genomic DNA"
             /strain="CLIB 89"
             /db_xref="taxon:4952"
             /clone="AM0A009H09"
             /clone_11b="AM0AA"
             /note="end : T7"

ORIGIN
Query Match      5.2%; Score 60; DB 9; Length 759;
Best Local Similarity 41.0%; Pred. No. 0.006;
Matches 123; Conservative 36; Mismatches 141; Indels 0; Gaps 0;

660 TGTGACCTTGTTCCGTTGTTTTTTTAAATTTTTCATAGTCTTTGTTTATCTTC 719

```

	Best Local Similarity	46.2%; Pred. No. 0.0077;	Matches	108; Conservative	14; Mismatches	112; Indels	0; Gaps	0;
OY	660	TGTCACCTTGTTCTCGTGGTTTTTTTAATTTTCATAAAGTCTTTGTATATCTC	719					
Db	263	TTTCTTAATTT	204					
OY	720	AATACAAATTTTGGCGCTGATNCTGCAAACTCTGCATCATATGCCCAATATAGTGAAC	779					
Db	203	TTTTTAATTTTTTGGTTTTTTTTTTTTTAATMTTITTTTTTTNNNNNNNGCNTNTNN	144					
OY	780	ACTGATCATCAAATTTGTGTGTTAATTGTTAAATTCATATCTCCGGTTTAAAA	839					
Db	143	NTTGGTGGGTATATATRTGTTATTTGTTTTTACTGNNTTTTTTTTTTTTTTTTTTT	84					
OY	840	GTCATTAATATGATCATGCGTTAAAACATTGTAAGAATGATTAATATAATGA	893					
Db	83	TTGAMTTTTTTTTTTTAAAGGTGAARTKWTAAAAAANDBBKADNNNNNRD	30					
RESULT 13								
CL648142		1187 bp DNA linear GSS 06-JUL-2000						
LOCUS	CH213-164G18.SP6 CH213	Gasterosteus aculeatus genomic clone						
DEFINITION	CH213-164G18 3', genomic survey sequence.							
ACCESSION	CU648142							
VERSION	CU648142.1 GI:49667566							
SOURCE	GSS.							
KEYWORDS	Gasterosteus aculeatus (three spined stickleback)							
ORGANISM	Gasterosteus aculeatus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Perciformes; Gasterosteidae; Gasterosteinae; Gasterosteus. 1 (bases 1 to 1187) Kingley, D., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M. Expressed sequence tags from Gasterosteus aculeatus Unpublished (2004) Contact: Grimwood, Jane Stanford Human Genome Center Stanford University School of Medicine 975 S California Avenue, Palo Alto, CA 94304, USA Tel.: 650 320 5917 Fax: 650 320 5801 Email: jane@shgc.stanford.edu Plate: 164 Class: BAC ends High quality sequence start: 12 High quality sequence stop: 664. Location/Qualifiers							
REFERENCE								
AUTHORS								
TITLE								
JOURNAL								
COMMENT								
FEATURES								
source								
/organism="Gasterosteus aculeatus"								
/mol_type="genomic DNA"								
/strain="Salmon River"								
/db_xref="taxon:69293"								
/clone="CH213-164G18"								
/sex="Mixed"								
/cell_type="Blood"								
/clone_1ib="CH213"								
/note="Vector: pTRAPAC2.1; Site_1: EcoRI; The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center, funded by the NIH Centers of Excellence in Genomic Science (CEGS) initiative (http://cegs.stanford.edu). The clone was isolated from the BAC library CHORI-213 built by Pieter deJongs in collaboration with the Stanford Genome Evolution Center (http://www.chori.org/bacpac/). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/ordering_information.htm)."								
ORIGIN								
Query Match	5.1%; Score 59.4; DB 9; Length 1187;							
Best Local Similarity	51.4%; Pred. No. 0.0086;							

Matches 132; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

QY 657 TCATGTGACCTGTTCCGTTTTTTTAAATTTTCATTAAGCTCTTTGTTATC 716
Db 896 TAAATTTTAAATGATTTATTTATTTATTTTAAATTTTGAATAATTTATTTATTTATTT 955
QY 717 TTCAATACAAATTTTGGCTGATCTTGAACCTTGTGATCATATGCCATATACGTG 776
Db 956 TTATTTTAAATTTTAAATATTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTT 1015
QY 777 AACATGCGATCAATTTGTTGTTATTTGTTAAATTTGATTTCTATTCGCGTTTA 836
Db 1016 AACTTTATGAAATATTTTGGTATTTTATTTTAAATTTTAAATTTTAAATTTTATTTT 1075
QY 837 AAAGTGAATTAATGATCATGTTTAAACATTTGTAAGTAAGATTAATAAATGATTA 896
Db 1076 TAAGATTTATTTAATATGATTTATTTATTTTAAATATGTTTATTAATATATAT 1135
QY 897 ATTTAGTTGATGATTA 913
Db 1136 TTNTATNTATTTATTTAA 1152

RESULT 14
CNS020K7/c 1092 bp DNA linear GSS 01-SEP-2000
LOCUS Tetracodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 222L11 of library G from Tetracodon nigroviridis, genomic survey
sequence.

ACCESSION AL175696
VERSION AL175696.1 GI:7813753
KEYWORDS GSS; genome survey sequence.
SOURCE Tetracodon nigroviridis
ORGANISM Tetracodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE 1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetracodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetracodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE 20359837
PUBMED 10899143

REFERENCE 3
(bases 1 to 1092)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetracodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetracodon.
Location/Qualifiers
1..1092
/organism="Tetracodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="222L11"
/clone_1fb="G"

FEATURES
source

ORIGIN /note="Genoscope sequence ID : COAG222CF06LP1-end : T7"

Query Match 5.1%; Score 59.2; DB 9; Length 1092;
Best Local Similarity 39.9%; Pred. No. 0.0095;
Matches 123; Conservative 47; Mismatches 137; Indels 1; Gaps 1;

QY 659 ATGTGTACCTGTTCCGTTTTTTTAAATTTTCATTAAGTCTTTGTTATCTT 718
Db 1025 ATATTTTATTTTAAATTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTT 966
QY 719 CAATACAAATTTTGGCTGATCTTGAACCTTGCATCATCGCAATATACGTGAA 778
Db 965 TTATATTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 907
QY 779 CACTGTGATCTAAATTTGTTGTTAAATTTGTTAAATTTGATTTCTATTCGCGTTTAA 838
Db 906 AAAAAAAMWMTTWTTTTATTTTATTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 847
QY 839 AGTGAATTAATGATCATGTTTAAACATTTGTAAGTAAGATTAATAAATGATTAAT 898
Db 846 AATATTTTWTTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 787
QY 899 TTAGTGAATGATTAACGGAACAAAAATGAGTATGATTTGTTGTTATTT 958
Db 786 WTAAMAMWMTTWTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 727
QY 959 TGACATAT 966
Db 726 TWAMTTT 719

RESULT 15
CNS03YE4/c 1045 bp DNA linear GSS 01-SEP-2000
LOCUS Tetracodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 068003 of library G from Tetracodon nigroviridis, genomic survey
sequence.

ACCESSION AL266197
VERSION AL266197.1 GI:7987962
KEYWORDS GSS; genome survey sequence.
SOURCE Tetracodon nigroviridis
ORGANISM Tetracodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.

REFERENCE 1
Roest Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetracodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL MEDLINE 20296633
PUBMED 10835645

REFERENCE 2
Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetracodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)

JOURNAL MEDLINE 20359837
PUBMED 10899143

REFERENCE 3
(bases 1 to 1045)
Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetracodon nigroviridis

COMMENT

genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

Location/Qualifiers

1..1045

/organism="Tetraodon nigroviridis"

/mol_type="genomic DNA"

/db_xref="taxon:99883"

/clone="068003"

/clone_lib="G"

/note="Genoscope sequence ID : C0BG068AH02LP2-end : T7"

ORIGIN

Query Match 5.1%; Score 58.8; DB 9; Length 1045;
Best Local Similarity 43.2%; Pred. No. 0.012;

Matches 115; Conservative 29; Mismatches 122; Indels 0; Gaps 0;

```
OY 669 TTGTTCCCTCGTTTTTTTAAATTTTCATTAAGTCTTTGTTTAACTTCAATACAAT 728
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1036 TTTTATTTTATTTTATTTTATTAATTTTAAATTTTAAATTTTATTTTATTAATA 977
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 729 TTTGGCTGATCTGCAAACTCTTCGATCATATGCGCAATATAGTGAACCTGGTAT 788
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 976 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 917
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 789 CTAATTTGTGTTAATTTGTAATTTAGATTCTATCTCCGTTTAAAGTGAATTAT 848
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 916 TTTATTTTWTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTWT 857
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 849 ATGTATCATGTTAAACATTTGTAAGTAAGATATATAATGATTAATTAGTTGATG 908
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 856 TTWMAATAAAMWMAAMWMAATTAAMWMTWTTTWTWTTTAAHAAARADARAR 797
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 909 GATAAGTGAGCAAAAATGAGATA 934
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 796 AATTTTATTAATAAATAAATAAATA 771
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: August 14, 2005, 07:20:46
Job time : 4428 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2005, 02:07:01 ; Search time 728 Seconds
(without alignments)
9416.292 Million cell updates/sec

Title: US-09-998-059-1

Perfect score: 1158
Sequence: 1 cacaacatcacatcaaat.....ctctcacaacaacaagaag 1158

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: geneseqn19808:*
2: geneseqn19908:*
3: geneseqn20008:*
4: geneseqn20018:*
5: geneseqn20028:*
6: geneseqn20038:*
7: geneseqn20048:*
8: geneseqn20058:*
9: geneseqn20068:*
10: geneseqn20078:*
11: geneseqn20088:*
12: geneseqn20098:*
13: geneseqn20108:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1158	100.0	1164	10	ADJ19001 Thale cre
2	1158	100.0	1223	6	ABZ17132 Arabidops
3	1064	91.9	1067	13	ADR89446 12S cruci
4	81	7.0	1118	3	ACG33222 Arabidops
5	65.4	5.6	10326	6	ABL34166 Human imm
6	64.2	5.5	547	13	ACN62109 Cotton gy
7	56.4	4.9	433	13	ACN51900 Cotton an
8	56.2	4.9	6000	6	ABL33133 Human imm
9	56.2	4.9	40324	6	ABO67149 Human ang
10	55.8	4.8	5845	6	ABL33662 Human imm
11	54.6	4.7	3476	13	AD889736 Oligonuc
12	54.6	4.7	7441	6	ABK40058 Human che
13	54.6	4.7	7479	6	AA653345 Chemical
14	54.2	4.7	6713	6	ABL33080 Human imm
15	54.2	4.7	6713	6	ABL34526 Human met
16	54.2	4.7	6713	6	ABL70251 Chemical
17	54.2	4.7	6713	7	AD899787 Bismulphit
18	53.6	4.6	7215	6	ABN91164 Staphyloc
19	53.6	4.6	7215	13	AD501346 Staphyloc
20	53.4	4.6	520	13	ACN52918 Cotton an

C	21	53.4	4.6	29993	10	ADB37662	Adb37662 Human che
	22	52.4	4.5	2000	8	ADA71938	Ada71938 Rice gene
	23	52.2	4.5	337	8	ABX41811	Abx41811 Bovine ES
	24	52	4.5	694	4	AAH53104	Aah53104 S. epider
	25	52	4.5	875	4	AAI95044	Aai95044 Human neu
	26	52	4.5	3028	4	AAH54496	Aah54496 S. epider
	27	52	4.5	3022	4	AAH54863	Aah54863 S. epider
	28	52	4.5	6968	4	AAH52683	Aah52683 S. epider
C	29	51.2	4.4	529	13	ACN56029	Adc56029 Cotton an
	30	51.2	4.4	96588	10	ADA03026	Ada03026 Human MBN
	31	51.2	4.4	96588	10	ADB72764	Adb72764 Human MBN
	32	51.2	4.4	96588	10	ADC85506	Adc85506 Human MBN
	33	51.2	4.4	96588	12	ADM74621	Adm74621 Human gar
C	34	51	4.4	11735	6	AA61141	Aa61141 Chemical
C	35	51	4.4	11735	6	ABK28167	Abk28167 DNA trans
C	36	51	4.4	11735	6	AA61141	Aa61141 Human gen
C	37	51	4.4	13511	6	ABL32280	Ab132280 Human imm
C	38	50.8	4.4	6621	8	ABZ10129	Abz10129 Haematopo
C	39	50.8	4.4	6621	8	ABZ10243	Abz10243 Haematopo
C	40	50.8	4.4	6621	10	ADE84153	Ad84153 Human lym
C	41	50.6	4.4	3683	8	ABZ10199	Abz10199 Haematopo
C	42	50.6	4.4	8011	6	AB132078	Ab132078 Human imm
	43	50.6	4.4	8011	6	ADP28368	Aad28368 Human che
	44	50.4	4.4	419	8	ABX46069	Abx46069 Bovine ES
	45	50.4	4.4	7544	4	AA645300	Aa645300 Chemical

ALIGNMENTS

RESULT 1	ADJ19001	standard; DNA; 1164 BP.
ID	ADJ19001	
XX	ADJ19001	
AC	ADJ19001	
XX	20-MAY-2004 (first entry)	
DT		
XX		
DE	Thale cress seed-specific promoter region Pl DNA.	
XX		
KW	plant; seed-specific promoter region; seed development; food;	
XX	agricultural; thale cress; ds; Pl.	
OS	Arabidopsis thaliana.	
XX		
PN	US2003005485-A1.	
XX	02-JAN-2003.	
PD		
XX		
PF	30-NOV-2001; 2001US-0098059.	
XX		
PR	01-DEC-2000; 2000US-0250401P.	
XX		
PA	(OHLE/) OHLEGE J B.	
PA	(BENN/) BENNING C.	
PA	(GAOH/) GAO H.	
PA	(GIRK/) GIRKE T A A.	
PA	(WHIT/) WHITE J A.	
XX		
PI	Ohlroge JB, Benning C, Gao H, Girke TAA, White JA;	
XX		
DR	WPI; 2003-370848/35.	
XX		
PT	New DNA comprising plant seed specific promoters, useful for controlling	
PT	expression of genes in plants, particularly for modifying seed products	
PT	(proteins, carbohydrates or oils), which are of major economic and food	
XX	values.	
XX		
PS	Claim 1: SEQ ID NO 1; 57bp; English.	
CC	The invention relates to a novel isolated DNA molecule which comprises a	
CC	plant seed-specific promoter region. The DNA molecule of the invention	
CC	may be useful for controlling the expression of genes in plants during	

the different phases of seed development, in particular, for modifying seed products, which are the major economic and food values of most agricultural plants e.g. as sources of proteins, carbohydrates or oils. The current sequence is that of the thale cress seed-specific promoter region P1 DNA of the invention.

XX Sequence 1164 BP, 399 A, 228 C, 205 G, 332 T, 0 U, 0 Other;

Query Match 100.0%; Score 1158; DB 10; Length 1164;
Best Local Similarity 100.0%; Pred. No. 5.2e-250;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CACAAACCTACCTAAATTCAGACTCACTCTAATTAATGCAATTCATCATGAA 60
Db 1 CACAAACCTACCTAAATTCAGACTCACTCTAATTAATGCAATTCATCATGAA 60
QY 61 AACATCAAAAAAGTCAGAAAGTAACAAATCAAGTGAATTCAGACACAAAGCCAGTAA 120
Db 61 AACATCAAAAAAGTCAGAAAGTAACAAATCAAGTGAATTCAGACACAAAGCCAGTAA 120
QY 121 GATAGAAAATTTTAAAGAAAGCTCATGTAGCTGGCAAAATTAATCTTAATCAAAACAG 180
Db 121 GATAGAAAATTTTAAAGAAAGCTCATGTAGCTGGCAAAATTAATCTTAATCAAAACAG 180
QY 181 TAAACAAGAGTAATTAGCAAAATCCAGACAGAAAATCTCACCCACCTCCGAATTCAG 240
Db 181 TAAACAAGAGTAATTAGCAAAATCCAGACAGAAAATCTCACCCACCTCCGAATTCAG 240
QY 241 TCTTCACCTAAATTTTGAAGAAGATGATCAATTCACCAATTAACCAAAATTCATTA 300
Db 241 TCTTCACCTAAATTTTGAAGAAGATGATCAATTCACCAATTAACCAAAATTCATTA 300
QY 301 TCAAAATGCGAGAAATCGTACCTGGAACCTTTCCTTCAAGTCGACAGAGAGAAAAGA 360
Db 301 TCAAAATGCGAGAAATCGTACCTGGAACCTTTCCTTCAAGTCGACAGAGAGAAAAGA 360
QY 361 AGATCGTGAAGAAAGGGTTTAGGTTTAACTCAGACTCTTATTTAGATTAATGGAGCG 420
Db 361 AGATCGTGAAGAAAGGGTTTAGGTTTAACTCAGACTCTTATTTAGATTAATGGAGCG 420
QY 421 GTGTCACTTTCCGTTTGGAAATGAACCTTGGGCTACGCTAATGGGCTTTAGATTT 480
Db 421 GTGTCACTTTCCGTTTGGAAATGAACCTTGGGCTACGCTAATGGGCTTTAGATTT 480
QY 481 TGATGGGCTTTCTAGTAATACAAATATTAAGTTATTTGGGCTTAATTAATAGCCCATGT 540
Db 481 TGATGGGCTTTCTAGTAATACAAATATTAAGTTATTTGGGCTTAATTAATAGCCCATGT 540
QY 541 TGAATATTTTGAACATGCTTGGCTTCTAGTCTTAAACATGCAACCGAACAGTTGTG 600
Db 541 TGAATATTTTGAACATGCTTGGCTTCTAGTCTTAAACATGCAACCGAACAGTTGTG 600
QY 601 AGACAAGTCGAGCAATATCAATGATCAACAGCCGTAAGTGGCCGCTCGCTAT 660
Db 601 AGACAAGTCGAGCAATATCAATGATCAACAGCCGTAAGTGGCCGCTCGCTAT 660
QY 661 GTGTCACTTTCCGTTTGGAAATGAACCTTGGGCTACGCTAATGGGCTTTAGATTT 720
Db 661 GTGTCACTTTCCGTTTGGAAATGAACCTTGGGCTACGCTAATGGGCTTTAGATTT 720
QY 721 ATACAAATTTTGGCTGATCTTGAACACTCTTCATCATTCGCAATATACGTAACA 780
Db 721 ATACAAATTTTGGCTGATCTTGAACACTCTTCATCATTCGCAATATACGTAACA 780
QY 781 CTGGGATCTAATTTTGTGTATATTTAATTAATTAATTAATTAATTAATTAATTAATTA 840
Db 781 CTGGGATCTAATTTTGTGTATATTTAATTAATTAATTAATTAATTAATTAATTAATTA 840
QY 841 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 900
Db 841 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 900
QY 901 AGTTGATGATTAACGTGAAGCAAAAAATGAGATGATGATGATTTGTCGATTTTG 960

```

```

Db 901 AGTTGATGATTAACGTGAAGCAAAAAATGAGATGATGATGATTTGTCGATTTTG 960
QY 961 ACATATGCGAGAGTAGAGCTACGCGCATGGAAGATCAAGAGACATTTGCTGAGCTCAAG 1020
Db 961 ACATATGCGAGAGTAGAGCTACGCGCATGGAAGATCAAGAGACATTTGCTGAGCTCAAG 1020
QY 1021 AGTACGTGTAAAAAGCTTAGAGTGAAGTCCCATGCAAACTTAATCTAGCTGGCTCAA 1080
Db 1021 AGTACGTGTAAAAAGCTTAGAGTGAAGTCCCATGCAAACTTAATCTAGCTGGCTCAA 1080
QY 1081 ACCAGAGCTCACTTGAACATATTAATCTCTTAAGTCCGTTCTTCAATCATCT 1140
Db 1081 ACCAGAGCTCACTTGAACATATTAATCTCTTAAGTCCGTTCTTCAATCATCT 1140
QY 1141 CTCACAAACAAACAAAAG 1158
Db 1141 CTCACAAACAAACAAAAG 1158

```

```

RESULT 2
AB217132
ID AB217132 standard; DNA; 1223 BP.
XX
AC AB217132;
XX
DT 21-JAN-2003 (first entry)
XX
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 4937.
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX
OS Arabidopsis thaliana.
XX
MO200216655-A2.
XX
PD 28-FEB-2002.
XX
PF 24-AUG-2001; 2001WO-US026685.
XX
PR 24-AUG-2000; 2000US-0227866P.
PR 26-JAN-2001; 2001US-0264647P.
PR 22-JUN-2001; 2001US-0300111P.
XX
PA (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI Harper JF, Kreps J, Wang X, Zhu T;
XX
WP1; 2002-304127/34.
XX
PT Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
PS Claim 144; SEQ ID NO 4937, 577bp + Sequence Listing; English.
XX
CC The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array of probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB21574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 1223 BP, 421 A, 241 C, 212 G, 349 T, 0 U, 0 Other;

```

Query Match 100.0%; Score 1158; DB 6; Length 1223;
Best Local Similarity 100.0%; Pred. No. 5.2e-250;

Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	CACAAACATACATCACTCAATTCATCTCAATTAAGCACTCATCA	60	
Db	63	CACAAACATACATCACTCAATTCATCTCAATTAAGCACTCATCA	122	
Qy	61	AACATCAAAACAGTCAAAAGTAAACAAATCAAGTCAAGTTCAGCACA	120	
Db	123	AACATCAAAACAGTCAAAAGTAAACAAATCAAGTTCAGCACA	182	
Qy	121	GATAGAAAATTTTACGAACGCTCATGCTAGTGGCCAAAATTTCTTAAT	180	
Db	183	GATAGAAAATTTTACGAACGCTCATGCTAGTGGCCAAAATTTCTTAAT	242	
Qy	181	TACAAACGAGTAATTAAGCAAAATCCGAGCAGAAAATCTCACCCACT	240	
Db	243	TACAAACGAGTAATTAAGCAAAATCCGAGCAGAAAATCTCACCCACT	302	
Qy	241	TCTTCACTAAAATTTTGAAGGAATCGATCAATACCAATTCACAAAATA	300	
Db	303	TCTTCACTAAAATTTTGAAGGAATCGATCAATACCAATTCACAAAATA	362	
Qy	301	TCAAAATGGCGAAGTCCGACCTCGAAACTTTGCTTCAAGTGGCAGAGAG	360	
Db	363	TCAAAATGGCGAAGTCCGACCTCGAAACTTTGCTTCAAGTGGCAGAGAG	422	
Qy	361	AGATCGTGGAGAAAGGGGTTTAGGTTAAGCTCAGACTTCTATTGAAT	420	
Db	423	AGATCGTGGAGAAAGGGGTTTAGGTTAAGCTCAGACTTCTATTGAAT	482	
Qy	421	GTGTCACTTTTCCGTTTGAAGATGAACCTTTGGGCTCAGTTATGGCT	480	
Db	483	GTGTCACTTTTCCGTTTGAAGATGAACCTTTGGGCTCAGTTATGGCT	542	
Qy	481	TGATGGGCTTTCTAGTAATTAACATATTAATTGGGCTTAGTTAAAT	540	
Db	543	TGATGGGCTTTCTAGTAATTAACATATTAATTGGGCTTAGTTAAAT	602	
Qy	541	TGGAATATTTGACATGCTTGGCTAGTGTAAACATGCAACGGAAGCT	600	
Db	603	TGGAATATTTGACATGCTTGGCTAGTGTAAACATGCAACGGAAGCT	662	
Qy	601	AGACAAGTCGACATATACAAATGATCAAAACGCGTATGCGCGTCTCAT	660	
Db	663	AGACAAGTCGACATATACAAATGATCAAAACGCGTATGCGCGTCTCAT	722	
Qy	661	GTGTCACTTTTCCGTTTGAAGATGAACCTTTGGGCTCAGTTATGGCT	720	
Db	723	GTGTCACTTTTCCGTTTGAAGATGAACCTTTGGGCTCAGTTATGGCT	782	
Qy	721	ATACAAATTTTGGGCTGATCTTGAACCTTTGATTCATTCGCAATATAC	780	
Db	783	ATACAAATTTTGGGCTGATCTTGAACCTTTGATTCATTCGCAATATAC	842	
Qy	781	CTGTGATCTAATTTGTTGTTAAATTTGTAATTTAGATTTCTCCGGTTT	840	
Db	843	CTGTGATCTAATTTGTTGTTAAATTTGTAATTTAGATTTCTCCGGTTT	902	
Qy	841	TGAATTAATATGATATCGTTTAAACATTTGTAAGTATTAATTAATTA	900	
Db	903	TGAATTAATATGATATCGTTTAAACATTTGTAAGTATTAATTAATTA	962	
Qy	901	AGTGAATGATTAACGTGAAGCAAAATTAAGATTAATTTGATTTTGG	960	
Db	963	AGTGAATGATTAACGTGAAGCAAAATTAAGATTAATTTGATTTTGG	1022	
Qy	961	ACATATGCGAGAGTGAAGTACGCGATGAAGTCAAGACACTTGTGAGT	1020	
Db	1023	ACATATGCGAGAGTGAAGTACGCGATGAAGTCAAGACACTTGTGAGT	1082	
Qy	1021	AGTGAATGATTAACGTGAAGCAAAATTAAGTATTAATTAATTAATTA	1080	
Db	1083	AGTGAATGATTAACGTGAAGCAAAATTAAGTATTAATTAATTAATTA	1142	

Qy	1081	ACCACAGGCTCACTTGACATATATTAACCTCTCTTAAGTCCGTTCTCTCAT	1140	
Db	1143	ACCACAGGCTCACTTGACATATATTAACCTCTCTTAAGTCCGTTCTCTCAT	1202	
Qy	1141	CTCACACAAACAAAAG 1158		
Db	1203	CTCACACAAACAAAAG 1220		
RESULT 3				
ADR89446				
ID ADR89446 standard; DNA; 1067 BP.				
XX				
AC ADR89446;				
XX				
DT 02-DEC-2004 (first entry)				
XX				
DE 12S cruciferin AtCrU3 promoter, PRO185.				
XX				
KW ss; fructose bi-phosphate aldolase; promoter; transcription terminator;				
KW transformation; transgenic; plant; rice; maize; wheat; barley; millet;				
KW oat; rye; sorghum; soybean; sunflower; canola; sugarcane; alfalfa; bean;				
KW pea; flax; lupinus; rapeseed; tobacco; tomato; potato; squash; papaya;				
KW poplar; cotton; Cyclin D2; Cyclin D3; rubisco activase;				
KW putative extensin; 12S cruciferin AtCrU3; FAD2 satutase; G3PDH-like.				
XX				
OS Arabidopsis thaliana.				
XX				
PN WO200407616-A2.				
XX				
PD 10-SEP-2004.				
XX				
PF 26-FEB-2004; 2004WO-EP050213.				
XX				
PR 27-FEB-2003; 2003EP-00075587.				
XX				
PA (CROP-) CROPDESIGN NV.				
XX				
PI Broekaert W, Hatzfeld Y;				
XX				
DR WPI; 2004-653388/63.				
XX				
PT New isolated Arabidopsis thaliana promoter capable of driving and/or				
PT regulating expression, useful in plant molecular biology, in particular				
PT for producing transgenic plants with desired traits.				
XX				
PS Claim 1; SEQ ID NO 6; 41bp; English.				
XX				
CC This sequence represents the 12S cruciferin AtCrU3 promoter, PRO185. This				
CC promoter sequence may be used in a genetic construct along with a				
CC heterologous nucleic acid sequence operably linked to the promoter, and				
CC optionally a 3' transcription terminator. This construct may be used to				
CC transform a host cell in the generation of a transgenic plant. The				
CC transgenic plant is rice, maize, wheat, barley, millet, oats, rye,				
CC sorghum, soybean, sunflower, canola, sugarcane, alfalfa, bean, pea, flax,				
CC lupinus, rapeseed, tobacco, tomato, potato, squash, papaya, poplar or				
CC cotton. The promoter sequences of the invention are useful in driving				
CC and/or regulating expression of an operably linked nucleic acid. The				
CC methods and compositions are also useful in the field of plant molecular				
CC biology, in particular for producing transgenic plants with desired				
CC traits and/or phenotypes.				
XX				
SQ Sequence 1067 BP; 353 A; 204 C; 198 G; 312 T; 0 U; 0 Other;				
Query Match 91.9%; Score 1064; DB 13; Length 1067;				
Best Local Similarity 100.0%; Pred. No. 6.3e-229;				
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	95	CAGATTGACGACACAAAGCGATTAAGTAATTTTACGAACGCTCATGCTA	154	
Db	1	CAGATTGACGACACAAAGCGATTAAGTAATTTTACGAACGCTCATGCTA	60	

QY 155 CGCAAAATACCTTCCTAATCAAAACAGTAACAACAGTAATTAGCAAAATCCGAGCAGAAA 214
Db 61 CGCAAAATACCTTCCTAATCAAAACAGTAACAACAGTAATTAGCAAAATCCGAGCAGAAA 120
QY 215 ACTCTCAGCCACCTCCGAAATTTACAGCTCTCACTCTAAATTTTCGAAAGAAATCGATCAT 274
Db 121 ACTCTCAGCCACCTCCGAAATTTACAGCTCTCACTCTAAATTTTCGAAAGAAATCGATCAT 180
QY 275 ACCAACCATTACACAAAATACATATCAAAATGGCGAAATCGTACCTGGAAACTTTGC 334
Db 181 ACCAACCATTACACAAAATACATATCAAAATGGCGAAATCGTACCTGGAAACTTTGC 240
QY 335 TTCAAGTGGCAGAGAGAGAGAGAGAGAGATCGTGAAGAAAGGGTTTAGGGTTAAAGTC 394
Db 241 TTCAAGTGGCAGAGAGAGAGAGAGAGAGATCGTGAAGAAAGGGTTTAGGGTTAAAGTC 300
QY 395 AGACTTCTATTGGAGTAATGGGACGGTGCACATTTCCGTTTGGAAATGAACCTTTGG 454
Db 301 AGACTTCTATTGGAGTAATGGGACGGTGCACATTTCCGTTTGGAAATGAACCTTTGG 360
QY 455 GCTCAGCTTATGGGCTATTAGATATTTGGATGGGCTTTCTAGTAATAACATATTAATTAT 514
Db 361 GCTCAGCTTATGGGCTATTAGATATTTGGATGGGCTTTCTAGTAATAACATATTAATTAT 420
QY 515 TGGGCTTAGTTAAATAAGCCCATGTTGAAATATTTGACACATGCTTGGCTACTAGTG 574
Db 421 TGGGCTTAGTTAAATAAGCCCATGTTGAAATATTTGACACATGCTTGGCTACTAGTG 480
QY 575 CTAAACATGACACGAAACAGTGTGCGAGACAAAGTGGCAATATGATGATCAACAC 634
Db 481 CTAAACATGACACGAAACAGTGTGCGAGACAAAGTGGCAATATGATGATCAACAC 540
QY 635 GCCTAGTGTCCCGGCTCGCTCATGTGTCACCTTGTTCCTCGTTTCTTTTAAATTT 694
Db 541 GCCTAGTGTCCCGGCTCGCTCATGTGTCACCTTGTTCCTCGTTTCTTTTAAATTT 600
QY 695 TCATTAAGTCTTTTGTTTTATCTTCAATCAAAATTTGGCTGATCTTGGCAACTCTTC 754
Db 601 TCATTAAGTCTTTTGTTTTATCTTCAATCAAAATTTGGCTGATCTTGGCAACTCTTC 660
QY 755 GATCATATGCGCAATATACGTGAACACTGGATCTAATTTGTGTATTTGTTAAT 814
Db 661 GATCATATGCGCAATATACGTGAACACTGGATCTAATTTGTGTATTTGTTAAT 720
QY 815 TTAAATTTCTATTCCTCGGTTTAAAGTGAATTAATGATATGATGTTAAACATTGTAA 874
Db 721 TTAAATTTCTATTCCTCGGTTTAAAGTGAATTAATGATATGATGTTAAACATTGTAA 780
QY 875 TTAAGTGTATTAATAATGATTAATTTAGTGAATGATGATGATGATGATGATGATGATG 934
Db 781 TTAAGTGTATTAATAATGATTAATTTAGTGAATGATGATGATGATGATGATGATGATG 840
QY 935 GATACATTTGATTTTGTGATTTTGAATATGCGAGAGTGAAGTACGCGCATGAAGAT 994
Db 841 GATACATTTGATTTTGTGATTTTGAATATGCGAGAGTGAAGTACGCGCATGAAGAT 900
QY 995 CAAGAGACACTTGTGCTGAGCTCACAGAGTGAAGTGAATAAAAGCTTAGAGTCCCA 1054
Db 901 CAAGAGACACTTGTGCTGAGCTCACAGAGTGAAGTGAATAAAAGCTTAGAGTCCCA 960
QY 1055 TGCAAACTTAATCTCACTGCTCAAAACAGAGCTCACTTGAACATATTAATCTCTC 1114
Db 961 TGCAAACTTAATCTCACTGCTCAAAACAGAGCTCACTTGAACATATTAATCTCTC 1020
QY 1115 CTAAAGTCCGTTCTCTCATGCTCTCAACAAACAAATAAG 1158
Db 1021 CTAAAGTCCGTTCTCTCATGCTCTCAACAAACAAATAAG 1064

RESULT 4
AAC33222/c
ID AAC33222 standard; DNA; 1118 BP.
XX

AC AAC33222,
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SFG ID NO: 2235.
DE Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
EN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.

PR	8-JUN-1999	99US-0139463P
PR	8-JUN-1999	99US-0139750P
PR	18-JUN-1999	99US-0139763P
PR	21-JUN-1999	99US-0139817P
PR	22-JUN-1999	99US-0139891P
PR	23-JUN-1999	99US-0140353P
PR	23-JUN-1999	99US-0140354P
PR	24-JUN-1999	99US-0140685P
PR	28-JUN-1999	99US-0140823P
PR	29-JUN-1999	99US-0140991P
PR	30-JUN-1999	99US-0141287P
PR	01-JUL-1999	99US-0141842P
PR	01-JUL-1999	99US-0142154P
PR	02-JUL-1999	99US-0142055P
PR	06-JUL-1999	99US-0142380P
PR	08-JUL-1999	99US-0142803P
PR	09-JUL-1999	99US-0142920P
PR	12-JUL-1999	99US-0142977P
PR	13-JUL-1999	99US-0143542P
PR	14-JUL-1999	99US-0143624P
PR	15-JUL-1999	99US-0144005P
PR	15-JUL-1999	99US-0144085P
PR	16-JUL-1999	99US-0144086P
PR	16-JUL-1999	99US-0144335P
PR	19-JUL-1999	99US-0144334P
PR	19-JUL-1999	99US-0144335P
PR	20-JUL-1999	99US-0144352P
PR	20-JUL-1999	99US-0144632P
PR	20-JUL-1999	99US-0144884P
PR	21-JUL-1999	99US-0144814P
PR	21-JUL-1999	99US-0145086P
PR	21-JUL-1999	99US-0145087P
PR	22-JUL-1999	99US-0145087P
PR	22-JUL-1999	99US-0145089P
PR	22-JUL-1999	99US-0145192P
PR	23-JUL-1999	99US-0145115P
PR	23-JUL-1999	99US-0145218P
PR	23-JUL-1999	99US-0145951P
PR	26-JUL-1999	99US-0145274P
PR	26-JUL-1999	99US-0145276P
PR	27-JUL-1999	99US-0145913P
PR	27-JUL-1999	99US-0145918P
PR	28-JUL-1999	99US-0145951P
PR	28-JUL-1999	99US-0145951P
PR	02-AUG-1999	99US-0146388P
PR	02-AUG-1999	99US-0146388P
PR	03-AUG-1999	99US-0146389P
PR	03-AUG-1999	99US-0147038P
PR	04-AUG-1999	99US-0147204P
PR	04-AUG-1999	99US-0147302P
PR	05-AUG-1999	99US-0147182P
PR	06-AUG-1999	99US-0147260P
PR	06-AUG-1999	99US-0147303P
PR	06-AUG-1999	99US-0147416P
PR	09-AUG-1999	99US-0147483P
PR	09-AUG-1999	99US-0147915P
PR	10-AUG-1999	99US-0148171P
PR	11-AUG-1999	99US-0148319P
PR	12-AUG-1999	99US-0148341P
PR	13-AUG-1999	99US-0148565P
PR	13-AUG-1999	99US-0148684P
PR	15-AUG-1999	99US-0149368P
PR	17-AUG-1999	99US-0149175P
PR	18-AUG-1999	99US-0149426P
PR	20-AUG-1999	99US-0149722P
PR	20-AUG-1999	99US-0149723P
PR	20-AUG-1999	99US-0149939P
PR	23-AUG-1999	99US-0149902P
PR	23-AUG-1999	99US-0149930P
PR	25-AUG-1999	99US-0150566P

PR	26-AUG-1999;	99US-01508842
PR	27-AUG-1999;	99US-01510652
PR	27-AUG-1999;	99US-01510662
PR	27-AUG-1999;	99US-01510800
PR	30-AUG-1999;	99US-01513030
PR	31-AUG-1999;	99US-01514382
PR	01-SEP-1999;	99US-01519300
PR	07-SEP-1999;	99US-01523630
PR	10-SEP-1999;	99US-01530700
PR	13-SEP-1999;	99US-01540182
PR	15-SEP-1999;	99US-01540392
PR	16-SEP-1999;	99US-01547792
PR	20-SEP-1999;	99US-01551392
PR	22-SEP-1999;	99US-01554682
PR	23-SEP-1999;	99US-01556592
PR	24-SEP-1999;	99US-01556482
PR	28-SEP-1999;	99US-01565662
PR	29-SEP-1999;	99US-01571172
PR	04-OCT-1999;	99US-01577532
PR	05-OCT-1999;	99US-01578652
PR	06-OCT-1999;	99US-01580292
PR	07-OCT-1999;	99US-01582322
PR	08-OCT-1999;	99US-01583232
PR	12-OCT-1999;	99US-01583692
PR	13-OCT-1999;	99US-01592232
PR	13-OCT-1999;	99US-01592242
PR	14-OCT-1999;	99US-01593392
PR	14-OCT-1999;	99US-01593302
PR	14-OCT-1999;	99US-01593312
PR	14-OCT-1999;	99US-01596372
PR	14-OCT-1999;	99US-01596382
PR	18-OCT-1999;	99US-01595842
PR	21-OCT-1999;	99US-01607412
PR	21-OCT-1999;	99US-01607672
PR	21-OCT-1999;	99US-01607702
PR	21-OCT-1999;	99US-01608142
PR	21-OCT-1999;	99US-01608152
PR	22-OCT-1999;	99US-01609802
PR	22-OCT-1999;	99US-01609812
PR	22-OCT-1999;	99US-01609862
PR	25-OCT-1999;	99US-01614042
PR	25-OCT-1999;	99US-01614052
PR	25-OCT-1999;	99US-01614062
PR	26-OCT-1999;	99US-01613592
PR	26-OCT-1999;	99US-01613612
PR	28-OCT-1999;	99US-01619202
PR	28-OCT-1999;	99US-01619932
PR	28-OCT-1999;	99US-01619932
PR	29-OCT-1999;	99US-01621422

Query Match	7.0%; Score 81; DB 3; Length 1118;
Best Local Similarity	100.0%; Pred. No. 2.5e-08;
Matches	81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	322 CTGAAACTTTGCTTCAAGTCGCAAGAGAGAAAGAAAGATCTCTGAGAAAGGGCTTT 381
Db	81 CTGAAACTTTGCTTCAAGTCGCAAGAGAGAGAAAGAAAGATCTCTGAGAAAGGGCTTT 22
QY	382 AGGGTTTAAGCTCAGACTTCT 402
Db	21 AGGGTTTAAGCTCAGACTTCT 1
RESULT 5	
ABL34166/c	
ID	ABL34166 standard; DNA; 10326 BP.
XX	
AC	ABL34166;
XX	
DT	26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 2139.
DE
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antihaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
de.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP007537.
XX
PR 30-JUN-2000; 2000DE-01032529.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PA (EPIC-) EPICENOMICS AG.
XX
PI Olek A. Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.
XX
PS Claim 1; SEQ ID NO 2139; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC leukaria degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention
CC
SQ Sequence 10326 BP; 2542 A; 104 C; 2689 G; 4991 T; 0 U; 0 Other;
Query Match 5.6%; Score 65.4; DB 6; Length 10326;
Best Local Similarity 48.8%; Pred. No.0.00013;
Matches 177; Conservative 0; Mismatches 186; Indels 0; Gaps 0;
QY 2 ACAACATACACTCAAAATCCAGACTCATCTACTCATTTATGCACTTCATGATGAA 61
DB 2764 ACCAAGTAACTCCAAAAAACCACATAAATTAATTAACCTTAACAAAAATTTCAACTAT 2765
QY 62 ACATCAAAAACAGTCAAAAGTAACTCAAGTTCAGCAGACAAAGCCGTAAG 121
DB 2704 TCACCTCAAAACACAAAAATTTTCTACTTAATAATTAACAAACAAACAAAAATTA 2645
QY 122 ATAGAAATTTAAAGAGCTCATGCTAAGCTGGCAAAATTAATCTCCAAATCAAAACGT 181
DB 2644 AAAAAATTTTAAACACCTTCGAAAAAATTAATTAACAAACGCAAACTTCACAAATTA 2585
QY 182 AACACAGATTAATTAAGCAAAATCCAGAGCAAAACTTCACCCCACTTCGAAATTCAGGT 241
DB 2584 TACATTCACATTTCTAAATTAATAAAACAATTTATCAAAACATTAATAAAAAATTACTA 2525
QY 242 CTTCACTAAATTTTTCGAAGGAATCGATCAATACAAACCATTAACAAATAATCATAT 301
DB 2524 TATACACCATCTTAATAAAAAAATAATCAATTAATAATTAATTCGAAAAATTAATTAAT 2465
QY 302 CAAATGGCGGAGATTCGTAAGTCTGGAATCTTTGCTTCAAGTCGCGAGAGAGAAAAAGAA 361
DB 2464 TAAAAATTAACAAAAATTTTAAACAAATTAATCTTAATCTCAAAATTAATAAAAAA 2405

QY 362 GAT 364
DB 2404 TAT 2402
RESULT 6
ACN62109/C
ID ACN62109 standard; cDNA; 547 BP.
XX
AC ACN62109;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton gynoecium tissue EST Clone ID: LIB3829-025-Q6-N6-C2, SEQ:16890.
XX
KW Cotton; plant; EST; expressed sequence tag; transgenic plant; gynoecium;
KW variety Nuotton33B; library LIB3829; molecular tag; molecular marker;
KW genetic mapping; molecular mapping; seed germination; plant growth;
KW plant quality; plant yield; plant breeding; tissue printing; ss.
OS Gossypium hirsutum.
XX
PN US2004123340-A1.
XX
PD 24-JUN-2004.
XX
PF 12-DEC-2001; 2001US-00021323.
XX
PR 14-DEC-2000; 2000US-0255619P.
XX
PA (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINC/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.
XX
PI Deikman J, Feng PCC, Fincher KU, Ziegler TE;
XX
DR WPI; 2004-479808/45.
XX
CC New isolated nucleic acid molecule that encodes a plant protein or its
CC fragment, useful for isolating a variety of agronomically significant
CC genes associated with plant growth, quality or yield, and as molecular
CC tags to map genes.
XX
PS Claim 1; SEQ ID NO 16890; 34pp; English.
XX
CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nuotton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determine whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a

CC cotton variety Nucleon33B gynoeceum tissue cDNA library (LIB3828). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at segdata.uspto.gov/sequence.html?docID=US20040123340
XX

Query Match 5.5%; Score 64.2; DB 13; Length 547;
Best Local Similarity 52.9%; Pred. No. 0.00012;
Matches 138; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 680 TTTTTCATTTTTCATTAAGTCTTTTGTTCATTCATACAAATTTTGGCTGA 739
DB 350 TTTTTCATTTTTCATTAAGTCTTTTGTTCATTCATACAAATTTTGGCTGA 791
QY 740 TCTTGCACCTCTGATCATATCGCCATATACGACGATGATCATTTGTG 799
DB 290 TTTTTCATTTTTCATTTTTCCTTTTTCCTTTTTCATTTTTCATTTTTCCT 231
QY 800 TGTTCATTTTTCATTTTTCATTTTTCCTTTTTCCTTTTTCATTTTTCCT 859
DB 230 TTTTTCATTTTTCATTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCCT 171
QY 860 TTTTTCATTTTTCATTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCCT 919
DB 170 TTTTTCATTTTTCATTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCCT 111
QY 920 GCAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 940
DB 110 AA 90

RESULT 7

ID ACN51900/C
AC ACN51900 standard; cDNA, 433 BP.

DT 02-DEC-2004 (first entry)

DE Cotton androecium tissue EST Clone ID: LIB3828-009-Q1-N6-D11, SEQ:6681.

KM Cotton; plant; EST; expressed sequence tag; transgenic plant; androecium;
KM variety Nucleon33B; library LIB3828; molecular tag; molecular marker;
KM genetic mapping; molecular mapping; seed germination; plant growth;
KM plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

PN US2004123340-A1.

PD 24-JUN-2004.

PF 12-DEC-2001; 2001US-00021323.

PR 14-DEC-2000; 2000US-0255619P.

PA (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FING/) FINCHER K L.
PA (ZIEG/) ZIEGLER T E.

PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;

XX WPI, 2004-479808/45.

PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.

PS Claim 1, SEQ ID NO 6681, 34bp; English.

CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN61099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoeceum
CC tissue, developing fibres, carpal walls and sepa from variety
CC Nucleon33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determine whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety Nucleon33B androecium tissue cDNA library (LIB3828). The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the US
CC patent office at segdata.uspto.gov/sequence.html?docID=US20040123340
XX

Query Match 4.9%; Score 56.4; DB 13; Length 433;
Best Local Similarity 50.0%; Pred. No. 0.0066;
Matches 141; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 657 TCATGTCACCTGTCCTGCTTTTTCATTTTTCATTAAGTCTTTTGTTCATC 716
DB 282 TTAATGTTTTCCTTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCATTAAT 223
QY 717 TTTTTCATTAAGTCTTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCATTA 776
DB 222 TTTTTCATTTTTCCTTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCATTA 163
QY 777 AACACCTGCTGATCTAATTTGTGTGTAATTTGTAATTTTATTCCTCGGTTA 836
DB 162 TTTTTCATTTTTCCTTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCATTA 103
QY 837 AAAGTGAATTATATGATCATGTTAAACATTTGATGATGATGATGATGATGAT 896
DB 102 TTTTTCATTTTTCCTTTTTCCTTTTTCCTTTTTCATTTTTCCTTTTTCATTA 43
QY 897 ATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 938
DB 42 AA 1

RESULT 8

ID ABL33133/C
AC ABL33133 standard; DNA, 6000 BP.

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1106.

KM Human; immune system disease; cytosine methylation; antiasthmatic;
KM neutrophilic; anti-HIV; anticonvulsant; ophthalmological;
KM antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;

```
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;  
XX ds.  
OS Homo sapiens.  
PN WO200200928-A2.  
PD 03-JAN-2002.  
PF 02-JUL-2001; 2001WO-EP007537.  
XX  
XX  
XX 30-JUN-2000; 2000DE-01032529.  
PR 01-SEP-2000; 2000DE-01043826.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
PA  
PI Olek A, Piepenbrock C, Berlin K,  
XX  
XX WPI; 2002-130909/17.  
DR  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful for  
PT diagnosis and treatment of diseases associated with abnormal cytosine  
PT methylation.  
XX  
XX Claim 1; SEQ ID NO 1106; 32pp + Sequence Listing; German.  
PS  
XX  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention  
XX  
XX Sequence 6000 BP; 1186 A; 274 C; 1753 G; 2787 T; 0 U; 0 Other;  
SQ  
Query Match 4.9%; Score 56.2; DB 6; Length 6000;  
Best Local Similarity 51.0%; Pred. No. 0.013;  
Matches 133; Conservative 0; Mismatches 128; Indels 0; Gaps 0;  
QY 40 ATTATGCACTTCATGATAAATCAAAACAGTCAAAATGACAAATCAATGACAGAT 99  
DB 4828 ATTAAAAACCCAAACCGTAAACCCCTAAACGACTTCAATTTAAAAACAACTCAAC 4769  
QY 100 TCAGACACAAAGCCAGTAAGATGAAAATTTAGCAAGCTCATGCTAAGCTGCGCA 159  
DB 4768 TCAACTCAACAACCTACAAAAAAATCTCTAAATATTCTCACTTCAACTCCCTCA 4709  
QY 160 AATACTTCTATATCAAAAACAGTAACACGAGTAATTAACAAAATCCGACGAAAACTCT 219  
DB 4708 ATAATCTACTAATTAATCTCAACCACTATATAAAAACAAAGCTAATCTAATCAACC 4649  
QY 220 CACCACCTCGGAATTCACGCTTCAATAAATTTGAAAAGAGATGATCATACCA 279  
DB 4648 TAAAAACACCAAAACCGTATCATTTCTAAAAATTAAACAACTACAAAAAACACAA 4589  
QY 280 CCCATTACAAAAATACATA 300  
DB 4588 AACATACTACAAAAAATATAA 4568
```

```
RESULT 9  
ABQ67149  
ID ABQ67149 standard; DNA; 40324 BP.  
XX  
XX ABQ67149;  
AC  
XX  
XX 28-AUG-2002 (first entry)  
DT  
XX  
DE Human angiogenesis associated polynucleotide SEQ ID NO 179.  
XX
```

```
KW Human; angiogenesis; methylation; eye disease; glaucoma; tumour;  
KM inflammation; rheumatoid arthritis; diabetic retinopathy; antidiabetic;  
XX macular degeneration; inflammatory bowel disease; Crohn's disease;  
KM antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
XX antiarteriosclerotic; ds.  
OS Homo sapiens.  
PN WO200246454-A2.  
PD 13-JUN-2002.  
PF 06-DEC-2001; 2001WO-EP014320.  
XX  
XX  
XX 06-DEC-2000; 2000DE-01061338.  
XX  
XX (EPIG-) EPIGENOMICS AG.  
PA  
PI Schacht O;  
XX  
XX WPI; 2002-500450/53.  
DR  
XX  
XX New nucleic acid fragments from chemically treated angiogenesis-  
PT associated genes, useful for determining methylation status, e.g. in  
PT diagnosis or treatment of cancer.  
XX  
XX Claim 1; SEQ ID NO 179; 41pp + Sequence Listing; German.  
PS  
XX  
XX The invention relates to a nucleic acid (I) comprising a segment of 18  
CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)  
CC having sequences (ABQ68971-ABQ67178) or their complements. (I), also  
CC related oligomers, are used to evaluate the methylation status and/or  
CC single-nucleotide polymorphisms, in angiogenesis-related genes, for  
CC diagnosis and treatment of eye diseases, proliferative retinopathy,  
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,  
CC diabetic retinopathy, macular degeneration caused by neovascularisation,  
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and  
CC Crohn's disease. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences  
SQ  
Sequence 40324 BP; 12279 A; 436 C; 8127 G; 19482 T; 0 U; 0 Other;  
Query Match 4.9%; Score 56.2; DB 6; Length 40324;  
Best Local Similarity 49.5%; Pred. No. 0.02;  
Matches 145; Conservative 0; Mismatches 148; Indels 0; Gaps 0;  
QY 669 TTGTTTCTCGTTTTTTTTTAAATTTTCATTAAGTCTTTGTTTATCTTCAATACAAAT 728  
DB 12605 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 12664  
QY 729 TTTTGGCTGTATCTTGCAAACTCTTCGATCATATCGCCATATACGTGAACATGTGTAT 788  
DB 12665 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 12724  
QY 789 CTAAATTTGTGTAAATGTTAAATTTAGATCTATTCCTCGGTTAAAAGTGAATAT 848  
DB 12725 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAATAAAT 12784  
QY 849 ATGTATCATGTTAAACATGTAAAGATGATAAATAAATAATTAATTAAGTGTG 908  
DB 12785 GTTAATTTAGTTTATTATTAAGTGTATTAATGAATAAATTAATTAATTAAGTGA 12844  
QY 909 GATAACGTGAAGCAAAAAATGAGATAGATACATTTGTTGCTATTGGA 961  
DB 12845 ATAAATGGGTAGTATATAGTATTAATGAATAAATTTTATTAATGAATTTT 12897
```

```
RESULT 10  
ABL33662  
ID ABL33662 standard; DNA; 5845 BP.  
XX  
XX ABL33662;  
AC
```

XX 26-MAR-2002 (first entry)
DT
XX
DE Human immune system associated gene SEQ ID NO: 1635.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
KM antiarteriosclerotic; anti-anaemic; cytosolic; noctropic;
KM neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KM antirheumatic; antirheumatic; antidiabetic; antipsoriatic;
KM antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KM acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KM neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
de.

XX Homo sapiens.
OS
XX
XX WO200200928-A2.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 02-JUL-2001; 2001WO-EP007537.
PF
XX
XX 30-JUN-2000; 2000DE-01032529.
PR
XX
XX 01-SEP-2000; 2000DE-01043826.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
XX WPI; 2002-130909/17.
DR

XX Nucleic acid comprising fragment of chemically modified gene, useful for
PT diagnosis and treatment of diseases associated with abnormal cytosine
PT methylation.

XX Claim 1; SEQ ID NO 1635; 32pp + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention

XX Sequence 5845 BP; 1651 A; 58 C; 1168 G; 2968 T; 0 U; 0 Other;

XX Query Match 4.8%; Score 55.8; DB 6; Length 5845;
Best Local Similarity 49.5%; Pred. No. 0.016;
Matches 144; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 669 TTGTTCTCGTTTTTAAATTTTCAATGTTCTTTGTTTATCTTCAATACAAT 728
DB 634 TTTTCTT 693
QY 729 TTTTGGCTGATCTTGCAACTCTTCGATCATATGCCCAATACGTGAACACTGTGAT 788
DB 694 TTTTCTT 753
QY 789 CTAAATTTGTTGTTAATTTGTTAATTTGATTTCTCCGGTTAAAGTGAATAT 848
DB 754 TTTTCTT 813
QY 849 ATGTATCATGTTAAACATTTGTAAGTAAGATATAATAAATGATATAATTTAGTTGATG 908
DB 814 AATTTTCTTTGATTAAGTTTGAATAATTTTATTAAGTTAATTTTCTTTTAAAT 873
QY 909 GATTAACGTGAACAAATAATGATATGATATCAATTTCTCCGTTAATTTT 959
DB 874 TGTATTTTAAATTAAGATTTAATTTTATTTTCTTTTTTTTTTTTTTTT 924

RESULT 11
ADS89736

ID ADS89736 standard; DNA; 3476 BP.

XX ADS89736;

AC 18-NOV-2004 (first entry)

DT Oligonucleotide of the invention SEQ ID NO:752.

XX ss; cell proliferative disorder; breast; methylation; cytosolic;
KM gene therapy; single nucleotide polymorphism; SNP.

XX Unidentified.

OS
XX
XX WO2004035803-A2.
PN
XX
XX 29-APR-2004.
PD
XX
XX 01-OCT-2003; 2003WO-EP010881.
PF
XX
XX 01-OCT-2002; 2002DE-01045779.
PR
XX
XX 07-JAN-2003; 2003DE-01000096.
PR
XX
XX 17-APR-2003; 2003DE-01017955.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Fockens J, Harbeck N, Koenig T, Maier S, Martens J, Model F;
PI
XX
XX Nimmrich I, Rujan T, Schmitt M, Look MP, Marx A;
PI
XX
XX WPI; 2004-348468/32.
DR

XX Predicting responsiveness of a subject with breast cell proliferative
PT disorder, useful for treating or differentiating breast cell
PT proliferative disorders comprises analyzing methylation pattern of a
PT genomic DNA from the subject.

XX Disclosure; SEQ ID NO 752; 104pp; English.

XX The invention relates to a novel method for predicting the responsiveness
CC of a subject with a cell proliferative disorder of the breast tissues to
CC a therapy comprising analysing the methylation pattern of a target
CC nucleic acid by contacting at least one of the target nucleic acids in a
CC biological sample obtained from the subject prior to or during treatment.
CC The method of the invention has cytosine activity, and may have a use
CC in gene therapy. The set of oligonucleotides comprising at least two of
CC the oligomers are useful for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms (SNPs) within the sequences. The
CC methods, nucleic acid, oligonucleotide, and kit are useful for the
CC treatment, characterisation, classification and/or differentiation, of
CC breast cell proliferative disorders. The method is also useful for
CC predicting the responsiveness of a subject with a cell proliferative
CC disorder of the breast tissues to a therapy. The present sequence is used
CC in the exemplification of the invention.

XX Sequence 3476 BP; 821 A; 0 C; 861 G; 1794 T; 0 U; 0 Other;

XX Query Match 4.7%; Score 54.6; DB 13; Length 3476;
Best Local Similarity 51.2%; Pred. No. 0.026;
Matches 152; Conservative 0; Mismatches 144; Indels 1; Gaps 1;

QY 669 TTGTTCTCGTTTTTAAATTTTCAATGTTCTTTGTTTATCTTCAATACAAT 728
DB 2114 TTTTGAATTTTTTTTATTTAGGAATTTTTTTTTTTTATTTTAAATTAATG 2173
QY 729 TTTTGGCTGATCTTGCAACTCTTCGATCATATGCCAATATACGTGAACACTGTGAT 788
DB 2174 AATGATATTTAATGAATAATTTGTTGATAGGAATAATGTTGAATATTAAGA 2233
QY 789 CTAAATTTGT-TGTGTTAATTTGTTAATTTGATTTCTCCGGTTAAAGTGAATTA 847
DB 2234 GTTGTATTTTATTTGTTAATTTTATTTAGTTATGATTAATTTTCTGTAATTAATG 2293

Oy	848	TATGATCATGGTTTAAACATTGTAGTAAGATGATTAATTAATGATTAATTTAGTTGAT	907
Db	2294	TTTTTTTGGATTATTAAGAAATTTAAAGATATTTAAAGTTAGGATTTTAAGTAAAGAAAT	2353
Oy	908	GGATAACGCGAAGCAAAAATGAGATAGATACATTGATTTTGTTCGATTTTGACAT	964
Db	2354	GGAGAAATGCTTGTAGATAGAGGAAAAATATGATTTTAAATTTTGTATTAATGTAGATAT	2410
	RESULT 12		
	ABK40058		
XX	ABK40058	standard; DNA; 7441 BP.	
AC	ABK40058,		
XX	21-MAY-2002	(first entry)	
DT			
XX			
DE		Human chemically pretreated gene sequence #70 strand 2.	
XX			
KW		Human; ds; bisulphite treatment; Cpg; DNA methylation; cancer; tumour;	
KW		cytosolic; ALDH6; CYP11A; CYP11B1; CYP3A5; DPYD; EPHX2; OCLN; TXNRD1;	
KW		UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.	
XX			
OS		Homo sapiens.	
XX			
PN	WO200202806-A2.		
PD	10-JAN-2002.		
XX			
PE	29-JUN-2001; 2001WO-EP007470.		
XX			
PR	30-JUN-2000; 2000DE-01032529.		
PR	01-SEP-2000; 2000DE-01043826.		
XX			
PA	(EPIC-) EPICGENOMICS AG.		
XX			
P1	Olek A, Piepenbrock C, Berlin K;		
DR	WPI; 2002-154757/20.		
PT		New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,	
XX		useful for detecting cytosine methylation state of genes associated with	
PT		pharmacogenomics and for therapy of diseases e.g. cancer.	
XX			
PS	Claim 1; SEQ ID NO 140; 24dp; English.		
XX			
CC		The invention relates to a nucleic acid comprising a sequence at least 18	
CC		bases in length of a segment of the chemically pretreated DNA of genes	
CC		associated with pharmacogenomics according to one of the sequences of the	
CC		genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1 (NM_000497), CYP3A3	
CC		(NM_000776 and NM_017460), DPYD (NM_000110), EPHX2 (NM_001979), OCLN	
CC		(NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360), MRP (NM_004996,	
CC		NM_019900, NM_019901, NM_019902, NM_019862, NM_019898, NM_019899) and	
CC		their complementary sequences, or a sequence (SI) chosen from 87	
CC		sequences and their complements. The chemical pretreatment is bisulphite	
CC		treatment to convert cytosines (but not methyl-cytosines) into uracils.	
CC		Also included are an oligomer (II) in particular an oligonucleotide or a	
CC		peptide nucleic acid (PNA)-oligomer, comprising in each case at least one	
CC		base sequence having a length of 9 nucleotides which hybridises to or is	
CC		identical to a chemically pretreated DNA of genes associated with	
CC		pharmacogenomics and their complements, arranged in an array for	
CC		analysing diseases associated with the methylation state (Cpg) and/or	
CC		detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The	
CC		oligomers may also be used as PCR primers. The set of 87 nucleic acids	
CC		and their complements is useful for diagnosis and therapy of solid	
CC		tumours and cancer. The present sequence represents one the 87 DNA	
CC		sequences or its complement. Note: The sequence data for this patent did	
CC		not form part of the printed specification, but was obtained in	
CC		electronic format directly from WIPO at	
XX		ftp.wipo.int/pub/published_pct_sequences	
XX			
SO	Sequence 7441 BP; 2128 A; 66 C; 1267 G; 3980 T; 0 U; 0 Other;		

[illegible]

DE Human metastasis associated gene SEQ ID NO: 79.

KX Metastasis associated gene; cytostatic; gene therapy; cancer;
KW cytosine methylation; gene; ds.
XX

OS Homo sapiens.

PN W020017376-A2.

PD 18-OCT-2001.

PF 06-APR-2001; 2001WO-EP003970.

PR 06-APR-2000; 2000DE-01019058.

PR 07-APR-2000; 2000DE-01019173.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

PA (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2002-010922/01.

XX New nucleic acid derived from chemically treated metastasis genes, useful
PT for diagnosis of cancers by analysis of cytosine methylation, also for
PT treatment.

PS Claim 1; SEQ ID NO 79; 23pp + Sequence Listing; English.

CC The present invention provides a number of human metastasis associated
CC genes which are modified by cytosine methylation. The sequences can be
CC used in the diagnosis and treatment of cancer. The present sequence is
CC one of the genes of the invention

XX
SQ Sequence 6713 BP; 1676 A; 149 C; 1441 G; 3447 T; 0 U; 0 Other;

Query Match 4.7%; Score 54.2; DB 6; Length 6713;
Best Local Similarity 49.1%; Pred. No. 0.038;
Matches 143; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 16 AAAATCCAGACTCATCTCACTTAATGCACTTCATCATGAAACATCAAAACAGT 75
DB 1045 AAAATATACCTCCACTCCATTAAATTAATTATATAAAAAACAAAACATTACCGA 986
QY 76 CAAAGTAACAAATCAAGTCAGATTCAGACACAAAGCCAGTAAAGATGAAATTTAAC 135
DB 985 TACAAAAATTAATCAAAAAAAACCCTCATACATCTAATAAATTAATTATAC 926
QY 136 GAACGCTCATGCTAAGCTGCGCAAAATCTTCCTAATCAAAAAGTAACACAGTAATT 195
DB 925 AACCACTATAAAAAACAATAATTCCTCAAAAAAATAAACTAATAACGATAATATA 866
QY 196 AGCAAAATCCGAGCAAAAACTCTCAGCCGACCTCCGAAATTACGCTTCTCACTAAATTT 255
DB 865 ATCTAACAAATCTTATATATATATATCTTAAAAAATAAAATCAATTAATTAATA 806
QY 256 TCGAAGAAGATGATCAATATACCAACCATTAACAAAAATATCAATCAAAA 306
DB 805 TCTACACTCCCATTTATTATTAACACATATTTCACAAATACTAAATATATA 755

Search completed: August 14, 2005, 04:36:46
Job time : 731 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 14, 2005, 02:10:26 ; Search time 5401 Seconds
(without alignments)
10389.028 Million cell updates/sec

Title: US-09-998-059-1
Sequence: 1 cacaacacacactcaaat.....ctctcacaacacacaaag 1158

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_ste:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1158	100.0	1158	6	AX707049 Sequence
2	1158	100.0	1158	6	AX707061 Sequence
3	1158	100.0	1233	6	AX510242 Sequence
4	1158	100.0	93655	8	ATP2009
5	1158	100.0	197655	8	ATCR169
6	1145.4	98.9	3054	8	ATU66594
7	1064	91.9	1067	6	CO875513 Sequence
8	956.2	82.6	3100	8	ATU66916
9	266	23.0	329	11	AX296428
10	198.8	17.2	4105	8	BNCRU1
11	193	16.7	6049	8	AF319771
12	181	15.6	4901	8	RSCRG
13	150	13.0	150	11	BX545778
14	119	10.3	180	11	BX296260
15	119	10.3	205	11	BX296270
16	81	7.0	1118	8	AY088047
17	65.4	5.6	10326	6	AX347068
18	62	5.4	1175	8	ATU66591
19	59.6	5.1	110000	8	CR382131_03

C	20	58	5.0	190639	5	BX572647	BX572647 Zebrafish
C	21	57.4	5.0	349751	3	PFMAL4P3	AL035476 Plasmodiu
C	22	56.6	4.9	7236	3	AF310890	AF310890 Dictyoste
C	23	56.4	4.9	110000	2	PFMAL13_11	Continuation (12 o
C	24	56.2	4.9	6000	6	AX346035	AX346035 Sequence
C	25	56.2	4.9	40324	6	AX458633	AX458633 Sequence
C	26	55.8	4.8	5845	6	AX346564	AX346564 Sequence
C	27	55.8	4.8	150101	5	BX248231	BX248231 Zebrafish
C	28	55.4	4.8	110000	3	AC116957_1	Continuation (2 of
C	29	55.4	4.8	349418	3	CR382398	CR382398 Plasmodiu
C	30	55.2	4.8	132549	2	CR392029	CR392029 Danio rer
C	31	55	4.7	153751	3	AC116551	AC116551 Dictyoste
C	32	54.8	4.7	257109	3	AC116963	AC116963 Dictyoste
C	33	54.6	4.7	3476	6	CO807302	CO807302 Sequence
C	34	54.6	4.7	7441	6	AX348445	AX348445 Sequence
C	35	54.6	4.7	7479	6	AX339173	AX339173 Sequence
C	36	54.4	4.7	33270	3	AC116921	AC116921 Dictyoste
C	37	54.4	4.7	337203	3	CR382401	CR382401 Plasmodiu
C	38	54.2	4.7	6713	6	AX281415	AX281415 Sequence
C	39	54.2	4.7	6713	6	AX345982	AX345982 Sequence
C	40	54.2	4.7	6713	6	AX348683	AX348683 Sequence
C	41	54.2	4.7	41399	3	AC116984_5	Continuation (6 of
C	42	54	4.7	1150	8	AY093228	AY093228 Arabidops
C	43	54	4.7	126845	5	AL954683	AL954683 Zebrafish
C	44	54	4.7	148385	2	CR749764	CR749764 Danio rer
C	45	53.8	4.6	273275	3	AE014828	AE014828 Plasmodiu

ALIGNMENTS

RESULT 1
AX707049 1158 bp DNA linear PAT 04-APR-2003
LOCUS AX707049
DEFINITION Sequence 1 from Patent WO03014347.
ACCESSION AX707049
VERSION AX707049.1 GI:29563360
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE
AUTHORS Ohlrogge, J. B., Benning, C., Gao, H., Gilke, T. A. and White, J. A.
TITLE Plant seed specific promoters
JOURNAL Patent: WO 03014347-A 1 20-FEB-2003;
MICHIGAN STATE UNIVERSITY (US)
FEATURES
source
1..1158
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 100.0%; Score 1158; DB 6; Length 1158;
Best Local Similarity 100.0%; Pred No. 4.5e-228;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAACATACACTCAAAATCCAGACTCAATCTACTCAATTATGCACTTCATGAA 60
DB 1 CACAAACATACACTCAAAATCCAGACTCAATCTACTCAATTATGCACTTCATGAA 60
QY 61 AACATCAAAACGCTAAAGTAAACAAATCAAGTCAATTCAGACACAAAGCAGTAA 120
DB 61 AACATCAAAACGCTAAAGTAAACAAATCAAGTCAATTCAGACACAAAGCAGTAA 120
QY 121 GATAGAAAATTTAAGCAAGCTCATGCTAAGCTGCGCAAAATCTCTATCAAAACAG 180
DB 121 GATAGAAAATTTAAGCAAGCTCATGCTAAGCTGCGCAAAATCTCTATCAAAACAG 180
QY 181 TAAACAAGGTAATTAAGCAAAATCCAGACGAGAAATCTTCACCCAGCTCGGAATTCAG 240

Db 181 TAACAACGAGTAATTAAGCAAAATCCGACGAGAAAATCTCACCACCTCCGAAATTCACG 240
Qy 241 TCTTCACATAAAATTTTCGAAAGAAATCGATCAATACCAACCCATTACACAAATATCATTA 300
Db 241 TCTTCACATAAAATTTTCGAAAGAAATCGATCAATACCAACCCATTACACAAATATCATTA 300
Qy 301 TCAAAATGGCGGAGATCGTACCTGGAACCTTTGCTTCAAGTGGCAGAGAGAGAGAAAAGA 360
Db 301 TCAAAATGGCGGAGATCGTACCTGGAACCTTTGCTTCAAGTGGCAGAGAGAGAGAAAAGA 360
Qy 361 AGATGCTGGAGAAAGGGGTTTAGGCTTAAAGCTCAGACTTCTATTGAGTAATAGGAGAG 420
Db 361 AGATGCTGGAGAAAGGGGTTTAGGCTTAAAGCTCAGACTTCTATTGAGTAATAGGAGAG 420
Qy 421 GTGTACACATTTCCGTTTGAAGATGAACTTTGGGCTCAGCTTATGGGCTATTAGATAT 480
Db 421 GTGTACACATTTCCGTTTGAAGATGAACTTTGGGCTCAGCTTATGGGCTATTAGATAT 480
Qy 481 TGATGGGCTTTCTAGTAATATCAATTAAGTTATGGGCTTATGGGCTTATAGATAT 540
Db 481 TGATGGGCTTTCTAGTAATATCAATTAAGTTATGGGCTTATGGGCTTATAGATAT 540
Qy 541 TCGAATATTTGACACATGCTTGGCTACTAGTGCCTAAACATGCAACCGAAGAGTTGTCG 600
Db 541 TCGAATATTTGACACATGCTTGGCTACTAGTGCCTAAACATGCAACCGAAGAGTTGTCG 600
Qy 601 AGACAAGTCGACGATATACAAATGATCAAAACGCTAGTGCCTGCGCTGCTCAT 660
Db 601 AGACAAGTCGACGATATACAAATGATCAAAACGCTAGTGCCTGCGCTGCTCAT 660
Qy 661 GTGTACCTTTGCTGCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 720
Db 661 GTGTACCTTTGCTGCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 720
Qy 721 ATACAATTTTGGCTGCTATCTTGCAAACTCTTCGATCATGCGCAATATACGTAACGTA 780
Db 721 ATACAATTTTGGCTGCTATCTTGCAAACTCTTCGATCATGCGCAATATACGTAACGTA 780
Qy 781 CTGGTATCTAATTTGTTGTGTTAATTTGTTAATTTGATTTCTATTTCTCGGTTTAAAG 840
Db 781 CTGGTATCTAATTTGTTGTGTTAATTTGTTAATTTGATTTCTATTTCTCGGTTTAAAG 840
Qy 841 TGAATTTATTTGATCATGTTTAAACATTTGTAAGTATGATTAATTAATTAATTAATTT 900
Db 841 TGAATTTATTTGATCATGTTTAAACATTTGTAAGTATGATTAATTAATTAATTAATTT 900
Qy 901 AGTTGATGATAACGTAAGCAAAATAGATAGATATGATTTGATTTTGTGTTATTTT 960
Db 901 AGTTGATGATAACGTAAGCAAAATAGATAGATATGATTTGATTTTGTGTTATTTT 960
Qy 961 ACATATGCGAGAGTAGCTACGCGCATGAAAGATCAAGAGACATTTGCTGAGCTCACAG 1020
Db 961 ACATATGCGAGAGTAGCTACGCGCATGAAAGATCAAGAGACATTTGCTGAGCTCACAG 1020
Qy 1021 AGTGAAGTGTAAAGGCTTGAAGTCCCATGCAAACTTATCTTACGTGGCTCAA 1080
Db 1021 AGTGAAGTGTAAAGGCTTGAAGTCCCATGCAAACTTATCTTACGTGGCTCAA 1080
Qy 1081 ACCAGAGCTCACTGACAAATATTAATCTCTCCCTAAGTCCGTTCTTATCATCAT 1140
Db 1081 ACCAGAGCTCACTGACAAATATTAATCTCTCCCTAAGTCCGTTCTTATCATCAT 1140
Qy 1141 CTCACACAAACAAAAG 1158
Db 1141 CTCACACAAACAAAAG 1158

RESULT 2
AX707061 1164 bp DNA linear PAT 04-APR-2003
LOCUS AX707061
DEFINITION Sequence 13 from Patent WO03014347.
ACCESSION AX707061
VERSION AX707061.1 GI:2956372

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Query Match 100.0%; Score 1158; DB 6; Length 1164;
Best Local Similarity 100.0%; Pred. No. 4.5e-228;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CACAAACATTAACCTCAAAATCCAGACTCAGATCTACTCAATTTATGCACTTCATCATGAA 60
1 CACAAACATTAACCTCAAAATCCAGACTCAGATCTACTCAATTTATGCACTTCATCATGAA 60
Qy 61 AACATCAAAAACAGTCAAAAGTACAAATCAAGTCAGATTGACGACACAAAGCCAGTAA 120
Db 61 AACATCAAAAACAGTCAAAAGTACAAATCAAGTCAGATTGACGACACAAAGCCAGTAA 120
Qy 121 GATGAAAAATTTTACGAACGCTCATGCTAGCTGCGCAAAATACTTCTTAATCAAAACAG 180
Db 121 GATGAAAAATTTTACGAACGCTCATGCTAGCTGCGCAAAATACTTCTTAATCAAAACAG 180
Qy 181 TAACAAGAGTAATTAAGCAAAATCCGACGAGAAAATCTCAACCCACCTCGAAATTCACG 240
Db 181 TAACAAGAGTAATTAAGCAAAATCCGACGAGAAAATCTCAACCCACCTCGAAATTCACG 240
Qy 241 TCTTCACATAAAATTTTCGAAAGAAATCGATCAATACCAACCCATTACACAAATATCATTA 300
Db 241 TCTTCACATAAAATTTTCGAAAGAAATCGATCAATACCAACCCATTACACAAATATCATTA 300
Qy 301 TCAAAATGGCGGAGATCGTACCTGGAACCTTTGCTTCAAGTGGCAGAGAGAGAGAAAAGA 360
Db 301 TCAAAATGGCGGAGATCGTACCTGGAACCTTTGCTTCAAGTGGCAGAGAGAGAGAAAAGA 360
Qy 361 AGATGCTGGAGAAAGGGGTTTAGGCTTAAAGCTCAGACTTCTATTGAGTAATAGGAGAG 420
Db 361 AGATGCTGGAGAAAGGGGTTTAGGCTTAAAGCTCAGACTTCTATTGAGTAATAGGAGAG 420
Qy 421 GTGTACATTTTCCGTTTGAAGATGAACTTTGGGCTCAGCTTATGGGCTATTAGATAT 480
Db 421 GTGTACATTTTCCGTTTGAAGATGAACTTTGGGCTCAGCTTATGGGCTATTAGATAT 480
Qy 481 TGATGGGCTTTCTAGTAATATCAATTAAGTTATGGGCTTATGGGCTTATAGATAT 540
Db 481 TGATGGGCTTTCTAGTAATATCAATTAAGTTATGGGCTTATGGGCTTATAGATAT 540
Qy 541 TCGAATATTTGACACATGCTTGGCTACTAGTGCCTAAACATGCAACCGAAGAGTTGTCG 600
Db 541 TCGAATATTTGACACATGCTTGGCTACTAGTGCCTAAACATGCAACCGAAGAGTTGTCG 600
Qy 601 AGACAAGTCGACGATATACAAATGATCAAAACGCTAGTGCCTGCGCTGCTCAT 660
Db 601 AGACAAGTCGACGATATACAAATGATCAAAACGCTAGTGCCTGCGCTGCTCAT 660
Qy 661 GTGTACCTTTGCTGCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 720
Db 661 GTGTACCTTTGCTGCTGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 720
Qy 721 ATACAATTTTGGCTGCTATCTTGCAAACTCTTCGATCATGCGCAATATACGTAACGTA 780
Db 721 ATACAATTTTGGCTGCTATCTTGCAAACTCTTCGATCATGCGCAATATACGTAACGTA 780

QY	781	CTGGTGAATCTAAATTGGTGTGATTAATTGATTCTAATTCGGGTTTAAAG	840
Db	781	CTGGTGAATCTAATTGTTGTTGCTTAATGTTAAATTAGATTCTAATTCGGGTTTAAAG	840
QY	841	TGAATTATATGATCATGGTTAAAAACATTGTAGTAAGATGATATAATAATGATAAATT	900
Db	841	TGAATTATATGATCATGGTTAAAAACATTGTAGTAAGATGATATAATAATGATAAATT	900
QY	901	AGTTGATGGAATAACGTCAAGCAAAAAATGAGATAGATACATTGTGATTTTGTCGATTTTG	960
Db	901	AGTTGATGGAATAACGTGAGCAAAAAATGAGATAGATACATTGTGATTTTGTCGATTTTG	960
QY	961	ACATATGCGGAGGTGAGCTACGGCGCATGAAGATCAAGAGACATGTCGAGCTCAAG	1020
Db	961	ACATATGCGGAGGTGAGCTACGGCGCATGAAGATCAAGAGACATGTCGAGCTCAAG	1020
QY	1021	AGTGCAGTGTAAAAAGCTTAGACTGAAGTCCCATGCAACCTAATCCTACGTGGCTCAA	1080
Db	1021	AGTGCAGTGTAAAAAGCTTAGACTGAAGTCCCATGCAACCTAATCCTACGTGGCTCAA	1080
QY	1081	ACCAAGAGCTACATTGACATATATTAAGTCTCTTAAGTCCGTTCTCTTCATCACTCT	1140
Db	1081	ACCAAGAGCTACATTGACATATATTAAGTCTCTCTTAAGTCCGTTCTCTTCATCACTCT	1140
QY	1141	CTCACAACAAACAAAAAG 1158	
Db	1141	CTCACAACAAACAAAAAG 1158	

RESULT 3			
AX510242			
LOCUS	1223 bp	DNA	linear
DEFINITION	AX510242		
	Sequence 4937 from Patent WO0216655.		PAT 27-SEP-2002

VERSION	AX510242.1	GI:23391479
KEYWORDS		
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	

ORGANISM
Arabiopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta, eudicotyledons, core eudicots,
rosids, eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1

REFERENCE
AUTHORS
TITLE
JOURNAL

1
Harper, J. F., Kreps, J., Wang, X. and Zhu, T.
Stress-regulated genes of plants, transgenic plants containing
same, and methods of use
Patent: WO 0216655-A 4937 28-FEB-2002;
US 2002/0012312 A1 2002-01-31

The Scripps Research Institute (US) ; Syngenta Participations AG (CH)

```
FEATURES
  source      Location/Qualifiers
1..1223
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
```

ORIGIN

Query Match	100.0%	Score 1158;	DB 6	Length 1223;
Best Local Similarity	100.0%	Pred. No. 4.5e-226;		
Matches 1158;	0;	Mismatches	0;	Gaps 0;
Conservative				

OY 1 CACAAACATACACCTCAAAATCCGAACTGACATCTCACTGAAATTATGCAACTTCATCATGAA 60
Db 63 CACAAACATACACCTCAAAATCCGAACTGACATCTCACTGAAATTATGCAACTTCATCATGAA 122

OY	61	AAACATCAAAAA	CAGTCAAA	GTAAACAAAAT	CAAGTCAGAT	TAGACA	CAACAAGCC	AGTAA	120
Db	123	AACATCAAAAA	CAGTCAAA	GTAAACAAAAT	CAAGTCAGAT	TAGACA	CAACAAGCC	AGTAA	182

121 GATAGAAATTTAAGCAGCGCTCATGCTAAGCTCGCAAAATCTTCTTAATCAAAACAG 180
 183 GATGAAATTTTAAAGCAGCGCTCATGCTAAGCTGGCAAAATATCTTCTTAATCAAAACAG 242

Dd 183 GATGAAATTTAA CGACGCTCATGCTAA GCTCGCAAAATACTTCTCTATCAAAACG 242

Oy 181 TAACCA CGAGTAATTAGCAA AATCCGAGCA GAAAACTCTCACCACCTCCGAAATTACG 240

Db 243 TAACAACGAGTAATTAGCAAAATCCGAGCAGAAAATCTCACCACCTCCGAATTACAC

Oy 241 TCTTCACAAATTTTTCGAAAGGATGATCATATCCAAACCCATTACCAAAATATCATTA 300
Db 303 TCTTCACAAATTTTTCGAAAGGATGATCATATCCAAACCCATTACCAAAATATCATTA 362

Qy 301 TCAAAATGGCGAATCTGTACCTGGAAACTTTGCTTCAAGTGTGCAGAGAGAGAGAAAAGGA 360

Db 363 TCAAAATGGCGAATCTGTACCTGGAAACTTTGCTTCAAGTGTGCAGAGAGAGAGAAAAGGA 422

Oy 361 AGATCGTGAGAAAGGGGTTTAGGTTAAAGCTCAGAGCTTCTATTGGAGTAAATGGGACG 430
 Db 423 AGATCGTGAGAAAGGGGTTTAGGTTAAAGCTCAGAGCTTCTATTGGAGTAAATGGGACG 482

421 GGTCCACATTTTCGTTTGGAAATGAACATTGGGCTACGGTTATGAGCTATATAGATAT 480
483 GTGTACACATTTTCGTTTGGAAATGAACATTGGGCTACGGTTATGAGCTATATAGATAT 542

Dy 481 TGAATGAGCTTTCAGTAATAACAAATATAGTTATGGCTTAGATTAAATAAGCCCATGT 540

Dd 543 TGATGGCCTTTCAGTAAATAACAATATTAAGTTAATGGCCTTAGTTAAATAAGCCCATGT 602

Oy 541 TGGAAATATTTGACACATGCTTGGCTACAGTGC TAAACATGCCAACGACACGT TGTG 800
 Db 603 TGGAAATATTTGACACATGCTTGGCTACAGTGC TAAACATGCCAACGACACGT TGTG 662

663 AGACAGTCGACGATATACCAATGATCAACACGCGCTAGTGTGCGCCGCTTCGCTCAT 722

Db 723 GGTGACCTTGTTTCGTTTTTTTAAATTTTCATPAGTTCCTTTGTTTATCTTCA 782

783 ATACAAATTTTGGCTGTATCTTGCAGAACTCTTGATCATATCGCCAAATATACGTGAACA 842

843 CCGCTGACTTAATTTGTTGTTGTTAATTTGTTAATTTGATTCTATTTCCGGTTTAAAG 902

Db

903 TCAATTATATGATCATGGTTAAACATGTTGTAAGTAAAGATGATTAATAAATGATAAATTT 962

963 AGTTGATGATTAACGTGAAGCAAAAAATGAGATAGATACATTGTGATTGTGCTATTTTG 10222

1023 ACATATCGCGAGAGTGAGCTACGCGCATGAAAGATCAAGAGACACTTGTCTCGAGCTCAAG 10822

D5	1083	AGTGACCTGMAAAAGCTTAGCTGAAGTCCCCATGCAGAACTTAATCTACGTGGCTCAA	11442
D6	1081	ACCACGAGCTCACTTGAACAATATATAAATCTCCCTAGTACCTCCGCTCTCATTCATCTCT	11440

Db	1143	ACCACGAGCTCACTTGACAATATATAAACTCCTCCTAAGTCCCGTTCTCTTCATCCATCT	1202
Dc	1141	CTCACAACAAAACAAAAAG	1158
Dy	1141	CTCACAACAAAACAAAAAG	1158

Db 1203 CTGACACAAACAAAAAG 1220

RESULT 4		
LOCUS	DEFINITION	
ATF2009	93695 bp	DNA linear PLN 03-AUG-1999
Arabidopsis thaliana DNA chromosome 4, BAC clone F2009 (ESSA project).		

ACCESSION	AL021749	GI:2842474
VERSION	AL021749.1	
KEYWORDS		
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid 11; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1	Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Hohenseil, J., Mewes, H. W., Mayer, K. F. X., Lemcke, K. and Schueller, C. Unpublished
AUTHORS	2 (bases 1 to 93695)	
JOURNAL	EU Arabidopsis sequencing project.	
REFERENCE	Direct Submission	
AUTHORS	Submitted (30-JUL-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: schueller@mpb.biochem.mpg.de, mayer@mps.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UD Norwich, UK, E-mail: michael.bevan@brc.ac.uk	
COMMENT	Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .	
FEATURES	Location/Qualifiers	
source	1. 93695	
	/organism="Arabidopsis thaliana"	
	/mol_type="genomic DNA"	
	/variety="Columbia"	
	/db_xref="taxon:3702"	
	/chromosome="4"	
gene	18. .467	
	/gene="F2009.10"	
	/complement (18. .467)	
gene	/gene="F2009.10"	
CDS	/complement (18. .467)	
	/gene="F2009.10"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="CA16872.1"	
	/db_xref="GI:2842475"	
	/db_xref="UniProt/TREMBL:O49442"	
exon	/translation="MTSYKLPFLVISKRTWACFLVYVAILSSSNLSINWYTG ENHQTSGLPAIVASVLGAVPGVLSMAALPIAIPAVIWIISVVTMAFGSRKR VVEGRKVTKEIGFVFRVLKEGNFVALLCALAVFVFENSYSSS"	
	/complement (18. .467)	
	/gene="F2009.10"	
	/number=1	
gene	2196. .2675	
	/gene="F2009.20"	
	/complement (2196. .2675)	
gene	/gene="F2009.20"	
CDS	/complement (2196. .2675)	
	/gene="F2009.20"	
	/codon_start=1	
	/product="hypothetical protein"	
	/protein_id="CA16873.2"	
	/db_xref="GI:5701784"	
	/db_xref="UniProt/TREMBL:O49443"	
	/translation="MADKEDEMTTSYKJFLKVIKRTWQCLFVVAVILSSRNS LNSIIVMGHNQTSGLPAIVASVLGAVPGVLSMAALPIAIPAVIWIISVVTI APFVKSRKVTKEIGFVFRVLKEGNFVALLCAVIAVFVFENSYSSS"	
exon	/complement (2196. .2675)	
	/gene="F2009.20"	
	/number=1	
gene	4143. .6188	
	/gene="F2009.40"	
CDS	/join(4143. .4941,4987. .5044,5096. .6188)	
	/gene="F2009.40"	
	/note="similarity to lectin receptor-like serine/threonine kinase IecRK1, Arabidopsis thaliana, PIR2:S68589	
	Contains protein kinases signatures and profile; pos. 328-349 & pos. 444-456"	

/codon_start=1	/product="receptor protein kinase like protein"
/protein_id="CA16875.1"	
/db_xref="GI:2842478"	
/db_xref="COA:049445"	
/db_xref="UniProt/TREMBL:O49445"	
/translation="MFESKVSILSLSLSLFRSTGTIEFIYNSNFTTTLILGNAT VSPSPITLTLNQTFPSIGRLYFSRIINASSSASPIPFATSTFISAPPKHLSPPHG FAFVLPFSETSAASSQHLFNFPTNNGDNSRIFAVEDVFNQGFNDINNHGV DVLNSITVASSETAGVGRDQRFTEKLNSGENYQAMIEFNGLAVNTARASSRKP IKPLISILNLTGLTLDLDMFGPTASQOLVQSHRILSMSEFNSNFIIGSVLSKGF IGVSSGVLLORLEGDVDEWETFEYMPRHVOYKDVLEATKGFSDENMIGYNSKVVVG VLEGEVAVKRIIMSPPRSVQATSEPLAEVNSIGRLAHKHVIGIKMSKSGSLLIY YEMNGSVYDKRIIPDCENLMEERMRVTRIDLASGMILYHEGMTKRIALHDISSNV LDKDNARVGDQFGLAKQLNTSKEMVSTTHVGTAGVNAPELVKTRGASQTDVYSFVG FVLVEVCGRRPIEEEREGIVMIGLMEKDKVVDGLDERIKANGVFVEVENALRIG ILICVHPDRVRPKNMGOVQILLEGRLVDDGERISILERKSSYLLETGEGSRQHP TRQDVYNNSSSYSNSQGTDSLHGR"	
4143. .4941	/gene="F2009.40"
/number=1	
4942. .4986	/gene="F2009.40"
/number=1	
4987. .5044	/gene="F2009.40"
/number=2	
5045. .5095	/gene="F2009.40"
/number=2	
5096. .6188	/gene="F2009.40"
/number=3	
6559. .11254	/gene="F2009.30"
/complement(join(6559. .6669,6878. .7050,7305. .7428,7537. .7647,7763. .7915,10629. .10733,10823. .10875,11074. .11254))	
/gene="F2009.30"	
/complement(join(6559. .6669,6878. .7050,7305. .7428,7537. .7647,7763. .7915,10629. .10733,10823. .10875,11074. .11254))	
/gene="F2009.30"	
/note="similarity to ENOD20 gene (small copper-binding protein), Medicago truncatula, PATCHX:E258656"	
/codon_start=1	
/product="copper-binding protein-like"	
/protein_id="CA16874.2"	
/db_xref="GI:5701785"	
/db_xref="UniProt/TREMBL:O49444"	
/translation="MGLVRFDLVLMFNLKGLFTISNGYKPYVGKDGWVPVPSD VSHGSHNRFOVNDTLPRQVPTFSVATLITVLELSSPPSPGVSESSNSNRVAP GPATPNSAGLSTSRDLQASBPVSSPLSSPALLDSGEBEKKTIIPKQKQAVLK LSIKQSPKQVNLVAALVRGMVEDALIQQLTVKRAAQTVVRHTTCFVNTLKKVXYI LQVIAARANAITHNGLDPRDLVIVIAEETPEEEAETARIKLVNFKKSKRREOLVP HTLIEITSPIVNRGRTKANHRSSSELVLTITLDVTCVMMEKVRDLNLTNQNNIYHNDE GP"	
/complement(6559. .6669)	
/gene="F2009.30"	
/number=1	
/complement(6670. .6877)	
/gene="F2009.30"	
/number=1	
/complement(6878. .7050)	
/gene="F2009.30"	
/number=2	
/complement(7051. .7304)	
/gene="F2009.30"	
/number=2	
/complement(7305. .7428)	
/gene="F2009.30"	
/number=3	

intron complement (7429. .7536)
/gene="F2009.30"
/number=3
exon complement (7537. .7647)
/gene="F2009.30"
/number=4
intron complement (7648. .7762)
/gene="F2009.30"
/number=4
exon complement (7763. .7915)
/gene="F2009.30"
/number=5
intron complement (7916. .10628)
/gene="F2009.30"
/number=5
cRNA 10056. .10127
/gene="F2009.30"
/product="tRNA-Lys"
/note="codon recognized: AAA"
exon complement (10629. .10733)
/gene="F2009.30"
/number=6
intron complement (10734. .10822)
/gene="F2009.30"
/number=6
exon complement (10823. .10875)
/gene="F2009.30"
/number=7
intron complement (10876. .11073)
/gene="F2009.30"
/number=7
exon complement (11074. .11254)
/gene="F2009.30"
/number=8
misc_feature 12220. .12639
/note="ESTmatch (T04156_F)"
13176. .16811
gene /gene="F2009.50"
join(13176. .13187,13379. .13519,13534. .13707,14167. .14233,
14308. .14388,14487. .14635,14753. .14783,15222. .15422,
15885. .16117,16178. .16213,16388. .16386,16629. .16739,
16804. .16811)
/gene="F2009.50"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA16876.2"
/db_xref="GI:5701786"
/db_xref="UniProt/TrEMBL:O49446"
/translation="MESKM,FGKQAEVGPFSANWITGYRGLPHALLHFYVYFYS
GISTMLNVLADYLSNVRKVSFSGTMRKLSLVNEMKHCNIEIAQVSRVASENNGD
TGVHYQLGIFSSQVFPESPRDLRKLSLVNEMKHCNIEIAQVSRVASENNGD

Query Match 100.0%; Score 1158; DB 8; Length 93695;
Best Local Similarity 100.0%; Pred. No. 2,8e-228;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAACATACATCAAAATCCAGACTCAGATCTACATCAATTTGAACCTTCATCTGA 60
Db 64001 CACAAACATACATCAAAATCCAGACTCAGATCTACATCAATTTGAACCTTCATCTGA 64060
QY 61 AACATCAAAAACAGTCAAAAGTAACAAATCAAGTCAAGTTCAGACACAAAGCCAGTAA 120
Db 64061 AACATCAAAAACAGTCAAAAGTAACAAATCAAGTCAAGTTCAGACACAAAGCCAGTAA 64120
QY 121 GATGAAAATTTTACGACGCTCATGCTAAAGCTGCCGAAAATCTTCTTAATCAAAAACAG 180
Db 64121 GATGAAAATTTTACGACGCTCATGCTAAAGCTGCCGAAAATCTTCTTAATCAAAAACAG 64180
QY 181 TAACAACGAGTATTTGCAAAATCCAGACGAAATCTTCACCCACTCCGAAATTCAGC 240
Db 64181 TAACAACGAGTATTTGCAAAATCCAGACGAAATCTTCACCCACTCCGAAATTCAGC 64240
QY 241 TCTTCACTAAATTTTGAAGAATCGATCAATACCAACCATTAACAAATATACATA 300

Db 64241 TCTTCACTAAATTTTGAAGAATCGATCAATACCAACCATTAACAAATATACATA 64300
QY 301 TCAAAATGGCGAGAAATCGATCGTGAACCTTTGCTCAAGTGGCAGAGAGAAAAAGCA 360
Db 64301 TCAAAATGGCGAGAAATCGATCGTGAACCTTTGCTCAAGTGGCAGAGAGAAAAAGCA 64360
QY 361 AGATCGTGGAGAAAGGGGTTTAGGTTTAAAGTCAAGCTTCTTAATTTGAAGTGAAGCA 420
Db 64361 AGATCGTGGAGAAAGGGGTTTAGGTTTAAAGTCAAGCTTCTTAATTTGAAGTGAAGCA 64420
QY 421 GTGTCAATTTTCCGTTTGGAAATGAACCTTTGGGCTCACGTTTGGGCTATTAGATT 480
Db 64421 GTGTCAATTTTCCGTTTGGAAATGAACCTTTGGGCTCACGTTTGGGCTATTAGATT 64480
QY 481 TGAATGGCTTTCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
Db 64481 TGAATGGCTTTCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 64540
QY 541 TGGAAATATTTTGACATGCTCTTGCTACTAGTCTAAACATGCAACCGAAGCTTGTG 600
Db 64541 TGGAAATATTTTGACATGCTCTTGCTACTAGTCTAAACATGCAACCGAAGCTTGTG 64600
QY 601 AGACAAAGTCGACGATATACATATGAATGAATCAACACGCTTACGCTGCTGCTCAT 660
Db 64601 AGACAAAGTCGACGATATACATATGAATGAATCAACACGCTTACGCTGCTGCTCAT 64660
QY 661 GTGTCACTTTGTTCTCGTTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 720
Db 64661 GTGTCACTTTGTTCTCGTTTCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 64720
QY 721 ATCAAAATTTTGGCTGTATCTTGCAAACTCTTGATCAATATGATGATTAATGATTAAT 780
Db 64721 ATCAAAATTTTGGCTGTATCTTGCAAACTCTTGATCAATATGATGATTAATGATTAAT 64780
QY 781 CTGTGATCTTAATTTGTTGTTAAATTTGTTAAATTTGTTAAATTTGTTAAATTTGTTAA 840
Db 64781 CTGTGATCTTAATTTGTTGTTAAATTTGTTAAATTTGTTAAATTTGTTAAATTTGTTAA 64840
QY 841 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
Db 64841 TGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 64900
QY 901 AGTTGATGATTAACGGAAGCAAAATGAATGAATGAATGAATGAATGAATGAATGAAT 960
Db 64901 AGTTGATGATTAACGGAAGCAAAATGAATGAATGAATGAATGAATGAATGAATGAAT 64960
QY 961 ACATATGCGAGAGTGAAGTACGCGCATGAATGAATGAATGAATGAATGAATGAATGAAT 1020
Db 64961 ACATATGCGAGAGTGAAGTACGCGCATGAATGAATGAATGAATGAATGAATGAATGAAT 65020
QY 1021 AGTACGTTTAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1080
Db 65021 AGTACGTTTAAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 65080
QY 1081 ACCAGAGCTCACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1140
Db 65081 ACCAGAGCTCACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 65140
QY 1141 CTCACAAACAAACAAAAG 1158
Db 65141 CTCACAAACAAACAAAAG 65158

RESULT 5
ATCHRIV69
LOCUS 197655 bp DNA linear PLN 16-MAR-2000
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 69.
ACCESSION AL161573.2 GI:7269705
VERSION
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 36863)	exon	/number=2 complement (6286. .6456) /gene="AT4g28510"
JOURNAL REFERENCE AUTHORS	Rose,M., Hempel,S., Entian,K.-D., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished 2 (bases 36754 to 80030)	intron	/number=3 complement (6457. .6574) /gene="AT4g28510"
JOURNAL REFERENCE AUTHORS	Lennard,N., Quail,M., Harris,B., Rajandream,M.A., Barrell,B.G., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished 3 (bases 79912 to 177066)	exon	/number=3 complement (6575. .6691) /gene="AT4g28510"
JOURNAL REFERENCE AUTHORS	Brandt,P., Dose,S., Jarke,D., Scharfe,M., Schon,O., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished 4 (bases 167625 to 197655)	intron	/number=4 complement (6692. .6935) /gene="AT4g28510"
JOURNAL REFERENCE AUTHORS	Hilbert,H., Braun,M., Holzer,E., Brandt,A., Duesterhoeft,A., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished 5 (bases 1 to 197655)	exon	/number=4 complement (6936. .7106) /gene="AT4g28510"
JOURNAL REFERENCE AUTHORS	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mps.biochem.mpg.de, mayer@mps.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/chal/ this fragment has an overlap with ATCHRIV68 at the 5' end and an overlap with ATCHRIV70 at the 3' end. Location/Qualifiers	gene	8330. .10351 /gene="AT4g28520"
FEATURES source	1. 197655 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4" 5704. .7106 /gene="AT4g28510" complement (join(5704. .5898, 5985. .6197, 6286. .6456, 6575. .6691, 6936. .7106)) /gene="AT4g28510" complement (join(5704. .5898, 5985. .6197, 6286. .6456, 6575. .6691, 6936. .7106)) /gene="AT4g28510" /note="strong similarity to B-cell receptor-associated protein BAP3 (prohibitin)-mouse contains EST gb:224518" /codon_start=1 /product="prohibitin-like protein" /protein_id="CAB81439.1" /db_xref="GI:7269706" /translation="NNNVKPKIPGGAIISTLTKGIIIGGLYCATSHLYNVEGGHR AIFNRVLGIXDKVYEGETHLMI PWFERPVIVDVARPYLVESTGSDLMQVKIGR VLTFRPMADLPETIYSGENYSERLPSIINETLKAAYQAQYASQLITQREAVSEIR KILTRRANFNVALDVISITNLTFRKEFTALBAEQVAQAEERAKFTVEAREQDKRS AVITRAQGEAKSQILIGQAIANNQAFITLRKLEAREIAQTIANSANKVYLSDDLLN LQGNMLDIVDAKN"	CDS	9444. .10351 /gene="AT4g28520"
exon	complement (5704. .5898) /gene="AT4g28510"	intron	/number=1 complement (5899. .5984) /gene="AT4g28510"
intron	complement (5899. .5984) /gene="AT4g28510"	gene	11459. .15337 /gene="AT4g28530"
exon	complement (5985. .6197) /gene="AT4g28510"	intron	complement (join(11459. .12034, 13178. .13332, 15121. .15337)) /gene="AT4g28530"
intron	complement (6198. .6285) /gene="AT4g28510"	CDS	complement (11459. .12034, 13178. .13332, 15121. .15337)) /note="similarity to No apical meristem gene (NAM), required for pattern formation in embryos and

flowers-*Pectunia hybrida*, PATCH: E205713
contains Aldehyde dehydrogenases active sites AA299-306
contains EST gb: R90477, R89959, A1992436.1, T04277,
T43017, R30004"
/codon_start=1
/product="NM / CUC2-like protein"
/protein_id="CA81441.1"
/db_xref="GI:7269708"
/db_xref="GOA:O49461"
/db_xref="UniProt/TREMBL:O49461"
/translation="MGLKDIKSLTPGPRFRHPSDELVCHYLCKNRKASDHQDVDD
DDVDLALKSTDLVEIDLHICPELPELVAKLNAKEMFESRDKVATGRTNAT
VSGWKATGKTRTMDPTROLQEDWVLCRVFNKRDSSLDNNYNNDDQOLEVD
ADPLANNOLPLSLSPPHNHQEKMKIQCVDQMDQKPSRTTGHPIYHCHQOTI
ACGDEOMIIGSLSPSSHPDHESLNLLVYDNNNSVNIISGDHQQVEKILLISLMT
SLDHDKTCMGSSSDGKMWSDIHWECGLSPETENILAPQ"
exon
/gene="AT4g28530"
complement(11459..12034)
/number=1
intron
complement(12035..13177)
/gene="AT4g28530"
/number=1
exon
complement(13178..13332)
/gene="AT4g28530"
/number=2
intron
complement(13333..15120)
/gene="AT4g28530"
/number=2
exon
complement(15121..15337)
/gene="AT4g28530"
/number=3
gene
28030..30462
/gene="AT4g28540"
join(28030..28105,28237..28277,28532..28601,28760..28908,
29004..29099,29338..29408,29503..29564,29655..29718,
29804..29888,29989..30115,30338..30462)
/gene="AT4g28540"
/note="strong similarity to protein kinase
ADK1-Arabidopsis thaliana
Contains Protein kinases signatures and profile
AA15-38; Protein kinases signatures and profile AA124-136
contains EST gb: F19756"
/codon_start=1
/product="protein kinase ADK1-like protein"
/protein_id="CA81442.1"
/db_xref="GI:7269709"
/db_xref="GOA:O65463"
Query Match 100.0%; Score 1158; DB 8; Length 197655;
Best Local Similarity 100.0%; Pred. No. 2,6e-228;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 7469 TCATAATGCGAGAACTGTAACCTGGAACCTTGCTTCAAGTCGACAGAGAGAGAAAAAGCA
Qy 361 AGATCGTGAGAAAGGGCTTTAGGGCTTTAAGCTGACACTTCTATTGAGTAATGGACG 420
Db 7529 AGATCGTGAGAAAGGGCTTTAGGGCTTTAAGCTGACACTTCTATTGAGTAATGGACG 7588
Qy 421 GGTGCACTTTCCGTTTGGAAATGAACTTTGGGCTCAGTTATGGGCTATTAGATTT 480
Db 7589 GTGTCACTTTCCGTTTGGAAATGAACTTTGGGCTCAGTTATGGGCTATTAGATTT 7648
Qy 481 TGAATGGCTTTCTAGTAATTAACAATTAATTAATTTGGGCTTAAATTAAGCCATGT 540
Db 7649 TGAATGGCTTTCTAGTAATTAACAATTAATTAATTTGGGCTTAAATTAAGCCATGT 7708
Qy 541 TGAATAATTTTGAACACATCTTGGCTACTAGTCTAAACATGCAACGAAAGCTGTG 600
Db 7709 TGAATAATTTTGAACACATCTTGGCTACTAGTCTAAACATGCAACGAAAGCTGTG 7768
Qy 601 AGACAAGTCGACGATATACAAATGATGATCAACACGCTAGTGGCGGCTCGCTCAT 660
Db 7769 AGACAAGTCGACGATATACAAATGATGATCAACACGCTAGTGGCGGCTCGCTCAT 7828
Qy 661 GTGTCACTTTGTTCTCGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTCA 720
Db 7829 GTGTCACTTTGTTCTCGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTCA 7888
Qy 721 ATACAAATTTTGGCTGCTGATCTTTGCAACTCTTCCATCTATGCGCAATTAAGTGAACA 780
Db 7889 ATACAAATTTTGGCTGCTGATCTTTGCAACTCTTCCATCTATGCGCAATTAAGTGAACA 7948
Qy 781 CTGTGATCTAATTTGTTGTTAATTTGTTAATTTGTTAATTTGTTAATTTGTTAATTTG 840
Db 7949 CTGTGATCTAATTTGTTGTTAATTTGTTAATTTGTTAATTTGTTAATTTGTTAATTTG 8008
Qy 841 TGAATTAATGATCATCTGTTTAAACATTTGTAAGTAAGTAATTAATTAATTAATTT 900
Db 8009 TGAATTAATGATCATCTGTTTAAACATTTGTAAGTAAGTAATTAATTAATTAATTT 8068
Qy 901 AGTTGATGATTAACGTAAGCAAAATTAAGATGATTAATTTGTTGTTGTTGTTGTTG 960
Db 8069 AGTTGATGATTAACGTAAGCAAAATTAAGATGATTAATTTGTTGTTGTTGTTGTTG 8128
Qy 961 ACATTAAGGAGAGAGAGTACGCGCATGAAGTCAAGAGACACTGCTGAGCTCACAG 1020
Db 8129 ACATTAAGGAGAGAGTACGCGCATGAAGTCAAGAGACACTGCTGAGCTCACAG 8188
Qy 1021 AGTGAAGTGAATAAAAGCTTAGAAGTGAAGTCCCATGCAAACTTAATCTAGTGTCA 1080
Db 8189 AGTGAAGTGAATAAAAGCTTAGAAGTGAAGTCCCATGCAAACTTAATCTAGTGTCA 8248
Qy 1081 ACCAGAGCTCTGACATTAATTAATCTCTTAAGTCCGTTCTTTATCCATCT 1140
Db 8249 ACCAGAGCTCTGACATTAATTAATCTCTTAAGTCCGTTCTTTATCCATCT 8308
Qy 1141 CTCACACACAAACAAAAG 1158
Db 8309 CTCACACACAAACAAAAG 8326

RESULT 6
ATU66594/c 3054 bp DNA linear PLN 05-JAN-1999
LOCUS Arabidopsis thaliana prohibitin 1 (Atphb1) gene, complete cds.
DEFINITION
ACCESSION U66594
VERSION U66594.1 GI:4097693
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 3054)

AUTHORS Sun, L. and Goodman, H.M.
TITLE Arabidopsis genes encoding prohibitin: importance for early development
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3054)
AUTHORS Sun, L., Nguyen, L. and Goodman, H.M.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1996) Molecular Biology Department, Massachusetts General Hospital, 50 Blossom St., Boston, MA 02114, USA
FEATURES
source 1..3054
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia"
/db_xref="taxon:3702"
987..3054
/gene="Atcphb1"
join(987..1048,1423..1602,1847..1963,2082..2252,2341..2553,2640..3054)
/gene="Atcphb1"
join(1432..1602,1847..1963,2082..2252,2341..2553,2640..2834)
/gene="Atcphb1"
/codon_start=1
/product="Prohibitin 1"
/protein_id="AAD00158.1"
/db_xref="GI:4097694"
/translation="MNNVVKPIPGGGAISTLLKVGIIIGLGLYGAATHSLVNEGGHR
AEMRLVIGDKVPEGTHLMIPWEEPVLYDVARPVLYVESTGSDLDLWIKIGLR
VUTRPMADQLBEIYRSLGSENVSRVLSINETLKAVYAOVRASLITQREAVSEIR
KILIERANFNVALDVSITNLTGKESTPALEAKQVAQERAKFTVEKAEODRS
AVIRAQGAQSAQLIGAIANNQAFITLRKIEAREIAQITANSANKVYLSDDLIN
LQGNMLDVDAKN"

ORIGIN

Query Match 98.9%; Score 1145.4; DB 8; Length 3054;
Best Local Similarity 99.8%; Pred. No. 1.6e-225;
Matches 1157; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CACAACATACACTCAAAATCCAGACTCAGATCTACTCAATATGCAACTTCATCATGAA 60
DB 1369 CACAACATACACTCAAAATCCAGACTCAGATCTACTCAATATGCAACTTCATCATGAA 1310
QY 61 AACATCAAAAACAGTCAAAAGTAACAAATCAAGTCAGATTCAGACACAAACCGAGTAAA 120
DB 1309 AACATCAAAAACAGTCAAAAGTAACAAATCAAGTCAGATTCAGACACAAACCGAGTAAA 1250
QY 121 GATTAAAAATTTAAGAAAGCTCATGCTAAGCTGGGCAAAATACCTTCTTAATCAAAACG 180
DB 1249 GATTAAAAATTTAAGAAAGCTCATGCTAAGCTGGGCAAAATACCTTCTTAATCAAAACG 1190
QY 181 TAACAACGATTAATAGCAAAATCCGAGCAAAAACCTCCACCCACTCCGAAATTTCAAG 240
DB 1189 TAACAACGATTAATAGCAAAATCCGAGCAAAAACCTCCACCCACTCCGAAATTTCAAG 1130
QY 241 TCTTCACTAAAAATTTGCGAAAGAAATCGATCAATACCAACCCATTACACAAATATCATPA 300
DB 1129 TCTTCACTAAAAATTTGCGAAAGAAATCGATCAATACCAACCCATTACACAAATATCATPA 1070
QY 301 TCAAAATGGCGAGATCGTAAGCTGCGAAATTTGCTTCAAGTGGCAGAGAGAGAAAGGA 360
DB 1069 TCAAAATGGCGAGATCGTAAGCTGCGAAATTTGCTTCAAGTGGCAGAGAGAGAAAGGA 1010
QY 361 AGATCGTGAAGAAAGGGTTTAGGTTAAGCTCAGACTCTATTGAGAGTAATGGAGAG 420
DB 1009 AGATCGTGAAGAAAGGGTTTAGGTTAAGCTCAGACTCTATTGAGAGTAATGGAGAG 950
QY 421 GTGTCAATTTTCGTTTGAAGATGAATTTGGGCTCAGTTATGGGCTATTAGATAT 480
DB 949 GTGTCAATTTTCGTTTGAAGATGAATTTGGGCTCAGTTATGGGCTATTAGATAT 890
QY 481 TGATGGGCTTTTATTAATATCAATATTAAGTATTGGGCTTAGTTAATTAAGCCATGT 540

DB 889 TGATGGGCTTTCTAGTAATAATCAATATTAAGTATTGGGCTTGTTTAAATAAGCCCATGT 830
QY 541 TGAATAATTTTGAACAATGCTCTGGCTACTAGTCTAAACATGCAACGGAACAGTTGTG 600
DB 829 TGAATAATTTTGAACAATGCTCTGGCTACTAGTCTAAACATGCAACGGAACAGTTGTG 770
QY 601 AGACAAGTCGACGATATATCAATGATGATCAACACGCGCTTAGTGTGCGCGCTCTGCAT 660
DB 769 AGACAAGTCGACGATATATCAATGATGATCAACACGCGCTTAGTGTGCGCGCTCTGCAT 710
QY 661 GTGTACCTGTGTTCCGCG-TTTTTTTTTAATTTTCAATAGTCTTTGTTTATCTC 719
DB 709 GTGTACCTGTGTTCCGCGTTTTTTTTTAATTTTCAATAGTCTTTGTTTATCTC 650
QY 720 AATCAAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATACGTGAAC 779
DB 649 AATCAAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATACGTGAAC 590
QY 780 ACTGTGATCTAATTTGTTGTTTAAATTTAGATTCTATTCTCCGCTTTAAA 839
DB 589 ACTGTGATCTAATTTGTTGTTTAAATTTAGATTCTATTCTCCGCTTTAAA 530
QY 840 GTGAATTTATGATCATGCTTAAACATTTGTAAGTAAAGATATTAATTAATTAAT 899
DB 529 GTGAATTTATGATCATGCTTAAACATTTGTAAGTAAAGATATTAATTAATTAAT 470
QY 900 TAGTTGATGATTAACGTGAGCAAAAATGAGATAGATACATTTGATTTGCTGATTTT 959
DB 469 TAGTTGATGATTAACGTGAGCAAAAATGAGATAGATACATTTGATTTGCTGATTTT 410
QY 960 GACATATCGGAGAGTGAAGTACCGGCAATGAAGATCAAGAGACCTTCTGAGCTCA 1019
DB 409 GACATATCGGAGAGTGAAGTACCGGCAATGAAGATCAAGAGACCTTCTGAGCTCA 350
QY 1020 GAGTACGCTTAAAAAGTTAGACTGAGAGTCCCATGCAAACTTAATCTACGTGCTCA 1079
DB 349 GAGTACGCTTAAAAAGTTAGACTGAGAGTCCCATGCAAACTTAATCTACGTGCTCA 290
QY 1080 AACCAAGCTCACTTGAACAATATTAATTAATCTCTTAAGTCCGCTTCTTCAATC 1139
DB 289 AACCAAGCTCACTTGAACAATATTAATTAATCTCTTAAGTCCGCTTCTTCAATC 230
QY 1140 TCTCAACAAACAAAAG 1158
DB 229 TCTCAACAAACAAAAG 211

RESULT 7
CQ875513 1067 bp DNA linear PAT 27-SEP-2004
LOCUS Sequence 6 from Patent WO2004076616.
DEFINITION CQ875513
ACCESSION CQ875513 GI:52748477
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (chale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 Broekaert, W. and Hatzfeld, Y.
AUTHORS Arabidopsis promoters
TITLE Patent: WO 2004076616-A 6 10-SEP-2004;
JOURNAL CropDesign N.V. (BB)
FEATURES
source 1..1067
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN

Query Match 91.9%; Score 1064; DB 6; Length 1067;
Best Local Similarity 100.0%; Pred. No. 9.9e-209;

OY	246	ACTATAAAATTTTCGAAAGAAATCGATTCATATACCAACCCCTTACAGAAAATACATATACAA	305
Db	71	ACTATAAAATTTTCGAAAGAAATCGATTCATATACCAACCCCTTACAGAAAATACATATACAA	130
OY	306	ATGGCGAAGATTCGATACCTCGAAGAACTTTTGCTTCAAGTCGACAGAGAGAGAAAGAGATC	365
Db	131	ATGGCGAAGATTCGATACCTCGAAGAACTTTTGCTTCAAGTCGACAGAGAGAGAAAGAGATC	190
OY	366	GTGGAGAAAGGGGTTTAGGGTTTAAAGTCAGACTTCTATTTGAGTTAAATGGAGCGGTCTC	425
Db	191	GTGGAGAAAGGGGTTTAGGGTTTAAAGTCAGACTTCTATTTGAGTTAAATGGAGCGGTCTC	250
OY	426	ACATTTTCCGTTTGGAAATGAACTTTGGGCTCACGTTATGGGCTATTAGTATTATTTGATG	485
Db	251	ACATTTTCCGTTTGGAAATGAACTTTGGGCTCACGTTATGGGCTATTAGTATTATTTGATG	310
OY	486	GGCTTTTCAGTAAATPACAATPAAAGTTATTTGGGCTTAGTTTAAATPAAAGCCATGTTGAA	545
Db	311	GGCTTTTCAGTAAATPACAATPAAAGTTATTTGGGCTTAGTTTAAATPAAAGCCATGTTGAA	370
OY	546	ATATTTCGACATGTCTTGGCTACTAGTCTAAACATGCAACGGAACGTTGTGAGACA	605
Db	371	ATATTTCGACATGTCTTGGCTACTAGTCTAAACATGCAACGGAACGTTGTGAGACA	430
OY	606	AGTGGCAGCATATPACATAGCATCAAAACGCGCTAGTGTGCGCGGCTCTCGCTCATGTGTC	665
Db	431	AGTGGCAGCATATPACATAGCATCAAAACGCGCTAGTGTGCGCGGCTCTCGCTCATGTGTC	490
OY	666	ACCTTGTTTCCTGCG-TTTTTTTTTTAAATTTTCAATPAGTCTTTTGTTTATCTTCAATAC	724
Db	491	ACCTTGTTTCCTGCGTTTTTTTTTTTATTTTCAATPAGTCTTTTGTTTATCTTCAATAC	550
OY	725	AAATTTTGGCTGATCTTGCAAACCTCTTGATCATATCGCCAAATPACGTGAACACTGG	784
Db	551	AAATTTTGGCTGATCTTGCAAACCTCTTGATCATATCGCCAAATPACGTGAACACTGG	610
OY	785	TGATCTAATTTTGTTGTGTTAATTTGTTAAATTTAATTTGATTTCTATTTCCCGSTTTAAAGTAA	844
Db	611	TGATCTAATTTTGTTGTGTTAATTTGTTAAATTTAATTTGATTTCTATTTCCCGSTTTAAAGTAA	670
OY	845	TTATATGATCATGGTTTAAACATTTGTAAGTAAAGATGATATPAAATGATPAAATTTAGTT	904
Db	671	TTATATGATCATGGTTTAAACATTTGTAAGTAAAGATGATATPAAATGATPAAATTTAGTT	730
OY	905	GATGATPACGTGAAGCAAAAAATGAGATPAGATCATTGTGATTTTGTGTAATTTTGCAT	964
Db	731	GATGATPACGTGAAGCAAAAAATGAGATPAGATCATTGTGATTTTGTGTAATTTTGCAT	790
OY	965	ATGCGGAGAGTGAAGTACCGCCATGTAAGATCAAGAGACATTTGCTCGAGCTCACAGAGTG	1024
Db	791	ATGCGGAGAGTGAAGTACCGCCATGTAAGATCAAGAGACATTTGCTCGAGCTCACAGAGTG	850
OY	1025	ACGTGTAAAAAGCTTAGACTGAAGTCCCATGCAAACTAATCTCTACGTGGCTCAAAACA	1084
Db	851	ACGTGTAAAAAGCTTAGACTGAAGTCCCATGCAAACTAATCTCTACGTGGCTCAAAACA	910
OY	1085	CGAGCTCACTTGACATATATPAACTCTCTCAATAGTCCGTTCTCTTCATCATCTCTCA	1144
Db	911	CGAGCTCACTTGACATATATPAACTCTCTCAATAGTCCGTTCTCTTCATCATCTCTCA	970
OY	1145	CAACCAAAACAAAAG 1158	
Db	971	CAACCAAAACAAAAG 984	

RESULT 9			
LOCUS	BX296428/c	329 bp	DNA
DEFINITION	Arabidopsis thaliana transposon insertion STS SW_3.33744, sequence		
ACCESSION	BX296428		
VERSION	BX296428.1	GI:29170295	

KEYWORDS	STIS: STS, sequence tagged site.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
REFERENCE AUTHORS	1 Clarke, J. H., Bowles, B., Carter, J., Hart, D., McCullagh, B., Murphy, G., Langham, S., LeGrys, C., Jones, J. D. G. and Bevan, M.
JOURNAL	Unpublished
REFERENCE AUTHORS	2 (bases 1 to 329)
TITLE	Clarke, J. H.
JOURNAL	Direct Submission
COMMENT	Submitted (24-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
COMMENT	AT denotes an activation tag dissociation transposon within a single line, Er an enhancer trap dissociation transposon, Gr a gene trap dissociation transposon, Mt a mis-expression enhancer trap dissociation transposon, Sm a defective suppressor mutator transposon. _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
COMMENT	On-line seed stock requests: http://nasc.nott.ac.uk/ NASC stock code: DEAD.
FEATURES	<p>Location/Qualifiers</p> <p>1..329</p> <p>/organism="Arabidopsis thaliana"</p> <p>/mol_type="genomic DNA"</p> <p>/variety="Columbia-0 NASC stock code N1092"</p> <p>/db_xref="taxon:3702"</p> <p>/clone="AL021749"</p> <p>/note="Derived from superpool 1.39 NASC code N40038"</p> <p>1..329</p> <p>/standard_name="SM_3.33744"</p>
ORIGIN	
STS	
Query Match	23.0%; Score 266; DB 11; Length 329;
Best Local Similarity	100.0%; Ptd. No. 1.6e-44;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	<p>893 ATAAATTAGTGATGAGTAAACGCGAAGCAAAAAATAGATAGATATGATTTGATTTGTC 952</p> <p>305 ATAAATTAGTGATGAGTAAACGCGAAGCAAAAAATAGATAGATATGATTTGATTTGTC 246</p> <p>953 GTATTTGACATATGCGAGAGTGAGCTACGCCGATGAATCAAGACACTTGCTCGA 1012</p> <p>245 GTATTTGACATATGCGAGAGTGAGCTACGCCGATGAATCAAGACACTTGCTCGA 186</p> <p>1013 GCTCACAAGAGTACGCTGTAAGCTTAAGCTGAAGTCCCATGCAAACTTAATCTTACG 1072</p> <p>185 GCTCACAAGAGTACGCTGTAAGCTTAAGCTGAAGTCCCATGCAAACTTAATCTTACG 126</p> <p>1073 TGGCTCAACACGAGGCTACCTGACAAATATATATATCTCTCCCTAAGTCCGCTCTTC 1132</p> <p>125 TGGCTCAACACGAGGCTACCTGACAAATATATATATCTCTCCCTAAGTCCGCTCTTC 66</p> <p>1133 ATCCATCTCTCACAACAACAAAAG 1158</p> <p>Db 65 ATCCATCTCTCACAACAACAAAAG 40</p>
RESULT 10	
LOCUS	BNCRU1 4105 bp DNA linear PLN 04-DEC-1993
DEFINITION	B.napus crul gene for cruciferin subunit.
ACCESSION	X62120 S48084
VERSION	X62120.1 GI:17800
KEYWORDS	12S storage protein; cruciferin.
SOURCE	Brassica napus (rape)
ORGANISM	Brassica napus
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.


```
/organism="Brassica napus"
/mol_type="genomic DNA"
/cultivar="Westar"
/db_xref="taxon:3708"
<54..>625
/product="CUC2-like protein"
<54..625
/codon_start=3
/product="CUC2-like protein"
/protein_id="AAK07611.1"
/db_xref="GI:12751304"
/translation="VNDGTVARFPHRPSKAYTTGTRKRRKRVHDSVDGSTRMKNKG
KTRPVLTGGRVGKYLIVLTNNGKOKKPKRTNMVMOYHLGTNEEKRELIYVSKY
FYQTPROCGGSAVAATAKAKRPYIHGSHHGGGGGHHLHYHLNHWKSGSAYAGE
YNNIPALISFNQGTIGQNLVNGSQPFP"
complement(join(<1106..1300,1388..1600,1686..1856,
1998..2114,2377..>2553))
/product="prohibitin 1-like protein"
complement(join(1106..1300,1388..1600,1686..1856,
1998..2114,2377..2553))
/codon_start=1
/product="prohibitin 1-like protein"
/protein_id="AAK07610.1"
/db_xref="GI:12751303"
/translation="MNLNNKVPKVPGGSAVSLKIGITGSLGYCATSLYVNDG
HRAIMFNKLVGKDYRPEGTMLVPMWERVYIDVRKRPVLVSTSGRLQWVKIG
LRLVTRPADQLPEIYRTLGNYSERVPSIITHETKAVQAQVAAQITQREAVSR
IKRLTORATFNIALDVSITTLTFGEFTPAIEAKQVAAQEAERAFIVEKEBOK
RSAYRAQGEAKSAQLIGQAIANNQAFITLKRKEAREIAQTIASHANKVYLSDDL
LNLOEMNLVNPVK"
3299..3317
/note="G-string"
3931..3936
join(<3991..4294,4411..4727,4824..5267,5528..>5932)
/product="cruciferin subunit"
join(3991..4294,4411..4727,4824..5267,5528..5932)
/codon_start=1
/product="cruciferin subunit"
/protein_id="AAK07609.1"
/db_xref="GI:12751302"
/translation="WVKLAHLVATPFGALVINGCLAROSLGVPPQIGNCLNDLIV
LOPRTTISEBAGRVYMDHNNPQIRACAVSRILIBGGGLYLPFFSPRISVYVG
MGISGRVVPGEFTPMDSQPMQOQGGQOQGGQOQGGQOQGGQOQGGQOQGGQOQ
EHRHGVIAITTAGSSHWITVNGDPLVITCLIDVLRQNDLDRKPIFRLAGNPNQ
GSQOQOQOQNMLSGFPQVLAQAKLIDVRLAQEQNDQSDRGKIVRVKGPQVVRP
LRPYESQOMHPRGPQSPQDNGLEETICSMRTHEINDDPAKDIVKPEVGRVTSV
SYTLPILOYLRLSATRGILQGNAMVLPKNNANMELICTGOARIQVNVNGQVLD
QOVKGOLVVIPOGRAVYVOSHONNFEMISRTNANMNVSTLAGRTSALRALPLEVLT
NAFOISLEBARIKNTLETTLTRARQPOLIEIYVA"
5953..5960
polya_signal
```

```
Db | 3627 ATAGCTTAATGTCAAAAAGTAGATTCCTATCCCTTTATATATAGCTTAAATGAAATT 3666
Qy | 856 ATGTTAAAAACATTGTAAGTAGATGATTAATAAATGATAAATTAGTTGATGATAACG 915
Db | 3687 TAAGCAGTTAAANATGAATTAATCA-ACAGAGAGTTCCTGTAAGTGTATTAACA 3745
Qy | 916 TGAACAAAAAATGAGATGATTAATGATTTGTGTGATTTTGGATTTGACATATGCGAGACT 975
Db | 3746 TGAACCTTAAATAAGAGATACA-TTTGAATTTGTGTGTTGTGGATATGCGAGATA 3804
Qy | 976 GAGCTAGCGCATGAAGATCAAGACACTGCTCGAGCTCAGACAGAGTAGAGTGAATAA 1035
Db | 3805 GTGCTAGCGCATGAACATCACTGAGACCTTGCTTCAGCTACAGAGTAGAGTGAATAA 3864
Qy | 1036 GCTTAGCTAGTCCCATGCAAACTTAATCTTACGTGCTCAAAACAGAGCTCACTT 1095
Db | 3865 CCATAGACCCACGACTTATGCAACCAAGCTTCACGCGGGAACCTTCATGCTCACTC 3924
Qy | 1096 GACAAATATATACTCTCTCTAG--TCCGTTCTTTGATCATCTCTCAACAA 1150
Db | 3925 CAC-ATATATAACTCTCAACCAAGCTTCATGTTCTTCATCATCTATCACCAAA 3980

RESULT 12
LOCUS RSCRUG 4901 bp DNA linear PLN 06-NOV-2003
DEFINITION Raphanus sativus pgcrure5 gene for cruciferin precursor.
ACCESSION X59808.4.7559
VERSION X59808.1 GI:21117
KEYWORDS 12S globulin; class II, cruciferin, seed storage protein.
SOURCE Raphanus sativus (radish)
ORGANISM Raphanus sativus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Raphanus.
1 (bases 1 to 4901)
REFERENCE Depigny-Thie,D., Raynal,M., Aspart,L., Delseny,M. and Grellet,F.
TITLE The cruciferin gene family in radish
JOURNAL Plant Mol. Biol. 20 (3), 467-479 (1992)
MEDLINE 93043037
PUBMED 1421150
REFERENCE 2 (bases 1 to 4901)
AUTHORS Depigny,D.M.C.
TITLE Direct Submission
JOURNAL Submitted (26-MAY-1991) Depigny D.M.C., University of Perpignan,
CNRS UA 565, Lab. de Physiologie Vegetale, Ave de Villeneuve, 66860
Perpignan, Cedex, FRANCE
On Nov 19, 2003 this sequence version replaced gi:258697.
COMMENT
FEATURES
source
location/Qualifiers
1..4901
/organism="Raphanus sativus"
/mol_type="genomic DNA"
/cultivar="Saxa Knacker"
/db_xref="taxon:3726"
/clone="Crure4"
678..685
704..2676
/gene="pgCrure5"
join(704..1043,1187..1470,1568..2014,2272..>2676)
/gene="pgCrure5"
/product="cruciferin precursor"
/evidence=experimental
704..1043
/gene="pgCrure5"
/number=1
join(740..1043,1187..1470,1568..2014,2272..2676)
/gene="pgCrure5"
/product="pgCrure5"
/codon_start=1
/product="cruciferin precursor"
/protein_id="CAA42478.1"
/db_xref="GI:21118"
/db_xref="GOA:Q02498"
```


Db 30 TCACAGAGTAGCGTGTAAAAAGCTTAGACT 1

RESULT 14
LOCUS BX296260/c
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.23228, sequence tagged site.
ACCESSION BX296260
VERSION BX296260.1 GI:29170125
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Murphy,G., Langham,S., Ledgrys,C., Jones,J.D.G. and Bevan,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 180)
AUTHORS Clarke,J.H.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
COMMENT Ar denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, Gf a gene trap dissociation transposon, SM a defective suppressor mutator dissociation transposon, _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon BBSRC GARNET, ATIS project
On-line seed stock requests: <http://nasc.nott.ac.uk/NASCstockcode:DEAD>.

FEATURES
source
1..180
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AU021749"
/note="Derived from superpool 16.17 NASC code N40712"
1..180
/standard_name="SM_3.23228"

ORIGIN
STS
Query Match 10.3%; Score 119; DB 11; Length 180;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 893 ATAAATTTAGTTGATGATAAGCGAAGCAAAAATGAGATGATGATTTGATTTGTC 952
DB 119 ATAAATTTAGTTGATGATAAGCGAAGCAAAAATGAGATGATGATTTGATTTGTC 60

QY 953 GTATTTTGACATATGCGGAGAGTGAGCTACCGCATGAAGATCAAGAGACACTTGCTCG 1011
DB 59 GTATTTTGACATATGCGGAGAGTGAGCTACCGCATGAAGATCAAGAGACACTTGCTCG 1

RESULT 15
LOCUS BX296270/c
DEFINITION Arabidopsis thaliana transposon insertion STS SM_3.23221, sequence tagged site.
ACCESSION BX296270
VERSION BX296270.1 GI:29170135
KEYWORDS STS; STS, sequence tagged site.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
AUTHORS Clarke,J.H., Bowles,B., Carter,J., Hart,D., McCullagh,B., Murphy,G., Langham,S., Ledgrys,C., Jones,J.D.G. and Bevan,M.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 205)
AUTHORS Clarke,J.H.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-2003) Clarke J.H., John Innes Centre, Colney Lane, Norwich, NR4 7UJ, UK
COMMENT Ar denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, Gf a gene trap dissociation transposon, SM a defective suppressor mutator dissociation transposon, _3 denotes a sequence derived from the 3' end of the transposon, _5 denotes a sequence derived from the 5' end of the transposon BBSRC GARNET, ATIS project
On-line seed stock requests: <http://nasc.nott.ac.uk/NASCstockcode:N110901>.

FEATURES
source
1..205
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AU021749"
/note="Derived from superpool 16.17 NASC code N40712"
1..205
/standard_name="SM_3.23221"

ORIGIN
STS
Query Match 10.3%; Score 119; DB 11; Length 205;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 893 ATAAATTTAGTTGATGATAAGCGAAGCAAAAATGAGATGATGATTTGATTTGTC 952
DB 119 ATAAATTTAGTTGATGATAAGCGAAGCAAAAATGAGATGATGATTTGATTTGTC 60

QY 953 GTATTTTGACATATGCGGAGAGTGAGCTACCGCATGAAGATCAAGAGACACTTGCTCG 1011
DB 59 GTATTTTGACATATGCGGAGAGTGAGCTACCGCATGAAGATCAAGAGACACTTGCTCG 1

Search completed: August 14, 2005, 06:06:56
Job time : 5406 secs

FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1601
LENGTH: 694
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-1601

Query Match 4.5%; Score 52; DB 4; Length 694;
Best Local Similarity 51.8%; Pred. No. 0.0013;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAGTCTCTTTGTTTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCTTGA 756
DB 246 ATTACCTCTGTTTATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 305
QY 757 TCATATCCCAATATACGGAACACTGGATCTAATTTGTGTGTTAATTGTTAAATTT 816
DB 306 TTCCACGCTACTATTTTATGAAGATCAATTAATCTCCATTTTACAACAATTTAAAT 365
QY 817 AGATTCTATTCCTCCGTTTAAAGTGAATTAATGATGATGATGATGATGATGATGATGAT 876
DB 366 TGCTTATACCCACTCTTTTAAAGTGAATATAGTATCTCGGATTTTATTAATTGGATTAA 425
QY 877 AGATGATATAATAATGATTAATTTAGTTGATGATGATGATGATGATGATGATGATGAT 924
DB 426 AGAATGATCAAAAGATGATCAAAATCAAGTTGTGTGTCATCAAA 473

RESULT 3
US-09-710-279-3860
Sequence 3860, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/164,258
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3860
LENGTH: 3028
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-3860

Query Match 4.5%; Score 52; DB 4; Length 3028;
Best Local Similarity 51.8%; Pred. No. 0.0022;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAGTCTCTTTGTTTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCTTGA 756
DB 2560 ATTACCTCTGTTTATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 2639
QY 757 TCATATCCCAATATACGGAACACTGGATCTAATTTGTGTGTTAATTGTTAAATTT 816
DB 2640 TTTCACGCTACTATTTTAAAGTGAATTAATGATGATGATGATGATGATGATGATGAT 2699
QY 817 AGATTCTATTCCTCCGTTTAAAGTGAATTAATGATGATGATGATGATGATGATGATGAT 876

DB 2700 TGCTTATACCCACTCTTTTAAAGTGAATATAGTATCTCGAATTTTATTAATTGGATTAA 2759
QY 877 AGATGATATAATAATGATTAATTTAGTTGATGATGATGATGATGATGATGATGATGAT 924
DB 2760 AGAATGATCAAAAGATGATCAAAATCAAGTTGTGTGTCATCAAA 2807

RESULT 4
US-09-710-279-4227
Sequence 4227, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4227
LENGTH: 3032
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-710-279-4227

Query Match 4.5%; Score 52; DB 4; Length 3032;
Best Local Similarity 51.8%; Pred. No. 0.0022;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAGTCTCTTTGTTTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCTTGA 756
DB 2689 ATTACCTCTGTTTATTTAAGATATATCAATTTCTAGATGGCGCTACCTACTTTTGA 2748
QY 757 TCATATCCCAATATACGGAACACTGGATCTAATTTGTGTGTTAATTGTTAAATTT 816
DB 2749 TTCCACGCTACTATTTTATGAAGATCAATTAATCTCCATTTTACAACAATTTAAAT 2808
QY 817 AGATTCTATTCCTCCGTTTAAAGTGAATTAATGATGATGATGATGATGATGATGATGAT 876
DB 2809 TGCTTATACCCACTCTTTTAAAGTGAATATAGTATCTCGGATTTTATTAATTGGATTAA 2868
QY 877 AGATGATATAATAATGATTAATTTAGTTGATGATGATGATGATGATGATGATGATGAT 924
DB 2869 AGAATGATCAAAAGATGATCAAAATCAAGTTGTGTGTCATCAAA 2916

RESULT 5
US-09-710-279-759
Sequence 759, Application US/09710279
Patent No. 6703492
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/09/710,279
CURRENT FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 759
LENGTH: 6968
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic

US-09-710-279-759

Query Match 4.5%; Score 52; DB 4; Length 6968;
Best Local Similarity 51.8%; Pred. No. 0.003;
Matches 118; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 697 ATAAAGTCTTTTGTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCTTCA 756
DB 246 ATTAAGTCTTTTGTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCTTCA 305
QY 757 TATATGCGCAATATACGGAACACTGCGATCTTAATTTGTGTATTTGAATTT 816
DB 306 TTTCAAGCTACTATTTTAAATGAAAGTCAATTTTCAATTTTCAATTTAATAT 365
QY 817 AGATCTATTTCCGGTTTAAAGTAAATATATATATATATATATATATATATATAT 876
DB 366 TGTATATACCACTCTTTTAAAGTAAATATATATATATATATATATATATATAT 425
QY 877 AGATGAT 924
DB 426 AGAAATGATCAAAAGATGATCAAAATCAAGTTGTGTCCATCAA 473

RESULT 6
US-08-894-731-2
Sequence 2, Application US/08894731
Patent No. 6084089

GENERAL INFORMATION:
APPLICANT: MINE, Toshiki
APPLICANT: OHYAMA, Akio
APPLICANT: HIYOSHI, Toru
APPLICANT: KASAKURA, Keisuke
TITLE OF INVENTION: COLD-INDUCIBLE PROMOTER SEQUENCE
FILE REFERENCE: 760-234P
CURRENT APPLICATION NUMBER: US/08/894, 731
CURRENT FILING DATE: 1997-10-27
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 2
LENGTH: 4140
TYPE: DNA
ORGANISM: Solanum tuberosum
US-08-894-731-2

Query Match 4.3%; Score 49.8; DB 3; Length 4140;
Best Local Similarity 48.1%; Pred. No. 0.009;
Matches 141; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 669 TTGTTCTCTGTTTCTTCTTCAATATAGTCTTTGTTTCTTCAATATAGTCTT 728
DB 1039 TTGTTCTCTGTTTCTTCTTCAATATAGTCTTTGTTTCTTCAATATAGTCTT 1098
QY 729 TTTTGGCTGATCTTCAAACTCTTCGATCATATGCGCAATATAGTCTTCAAT 788
DB 1099 TATCGTTGTTTAT 1158
QY 789 CTAATTTGTTGTTTATATATATATATATATATATATATATATATATATATAT 848
DB 1159 GTAAT 1218
QY 849 ATGTATCATGTTTAAATATATATATATATATATATATATATATATATATATAT 908
DB 1219 ATGTATCATGTTTAAATATATATATATATATATATATATATATATATATATAT 1278
QY 909 GAT 961
DB 1279 TACAATCTGAAAT 1331

RESULT 7
US-09-601-198-51
Sequence 51, Application US/09601198
Patent No. 6531583

GENERAL INFORMATION:
APPLICANT: Cassell, Gail H.
APPLICANT: Chen, Ellison Y.
APPLICANT: Glaes, Jennifer S.
APPLICANT: Glaes, John I.
APPLICANT: Heiner, Cheryl R.
APPLICANT: Lefkowitz, Elliot
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
FILE REFERENCE: UAB-13452/22

QY 693 TTTCAATAGTCTTTTGTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCT 752
DB 385 TGTATATACCTGTTTGGAGCAAGTTTAAATTTATATATATATATATATATATAT 444
QY 753 TCGATCATATGCGCAATATATATATATATATATATATATATATATATATATAT 812
DB 445 TTTCAATAGTCTTTTGGAGCAAGTTTAAATTTATATATATATATATATATATAT 504
QY 813 ATTTAGATCTTATCTCCGTTTAAAGTAAATATATATATATATATATATATATAT 872
DB 505 GTTTTCAAT 564
QY 873 AGTAAGAT 932
DB 565 ATTAATTTGTTTAAATTTGATATATATATATATATATATATATATATATATATAT 624
QY 933 TAGATCATTTGATTTTGTGATTTTGAATATATATATATATATATATATATAT 968
DB 625 TTGTGATATAGCTTTTGAATATATATATATATATATATATATATATATATAT 660

Query Match 4.2%; Score 48.8; DB 4; Length 1851;
Best Local Similarity 48.6%; Pred. No. 0.012;
Matches 134; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 693 TTTCAATAGTCTTTTGTATCTTCAATACAAATTTTGGCTGATCTTGCAACTCT 752
DB 385 TGTATATACCTGTTTGGAGCAAGTTTAAATTTATATATATATATATATATATAT 444
QY 753 TCGATCATATGCGCAATATATATATATATATATATATATATATATATATATAT 812
DB 445 TTTCAATAGTCTTTTGGAGCAAGTTTAAATTTATATATATATATATATATATAT 504
QY 813 ATTTAGATCTTATCTCCGTTTAAAGTAAATATATATATATATATATATATATAT 872
DB 505 GTTTTCAAT 564
QY 873 AGTAAGAT 932
DB 565 ATTAATTTGTTTAAATTTGATATATATATATATATATATATATATATATATATAT 624
QY 933 TAGATCATTTGATTTTGTGATTTTGAATATATATATATATATATATATATAT 968
DB 625 TTGTGATATAGCTTTTGAATATATATATATATATATATATATATATATATAT 660

RESULT 8
US-09-806-708B-22
Sequence 22, Application US/09806708B
Patent No. 6784342

GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806, 708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patent version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1)-(1141)

OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match 4.2%; Score 48.6; DB 4; Length 1141;
Best Local Similarity 10.2%; Pred. No. 0.012;

[illegible]

```

RESULT 13
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849) B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

```

	Query Match	4.0%;	Score 46;	DB 4;	Length 612;
	Best Local Similarity	51.2%;	Pred. No. 0.043;	Mismatches 101;	Indels 0; Gaps 0
	Matches	106;	Conservative	0;	
Qy	669	TTCGTTCTCGTTTTTTTTTAATTTTTCATAAGTCTTTGGTTTATCTTCATACAAAT	728		
Db	539				
		TT			480
Qy	729	TTTTGGCTGTACTGTCAACCTCTGCATCATATGCCAATTACGTGAACCTGTGAT	788		
Db	479				
		TTTATTTTTTTTTTTTTTTTATATATTTTATATTTTAAAATTTTTTTTTTTTTTTTTTTT			420
Qy	789	CTAAATTGTGTGAATGTGTAAATTTAGAATCTATTTCCGGTTTAAAAAGTGATAT	848		
Db	419	TTTTTATTTTTTTTTTTTTTTTATATTTTAAATTTTATTTTTTTTTTTTTTTTTTTTTT	360		
Qy	849	ATGTATCATGGTTAAACAATGTAGT	875		
Db	359	ATTTATTTTTTTTTTATTTTTTTAATT	333		

US-08-232-463-14/c
Sequence 14: Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-f15
US-08-232-463-14

	Query Match	4.0% ; Score 45.8 ; DB 1 ; Length 7218 ;
	Best Local Similarity 2.1% ; Pred. No. 0.12 ;	
	Matches 8 ; Conservative 217 ; Mismatches 154 ; Indels 0 ; Gaps 0 ;	
Oy	58 GAAAAATCATAAAACAGTCAAAGTAACAAAATCAGTCAGATTGACGACACAAAGCCAGT	117
Db	1431 RRR	1372
Oy	118 AAAATATGAAAATTTAACGAACGCTCATGTCACTGCGCAAAATACTCTTAATCAAAA	177
Db	1371 RRR	1312
Oy	178 CAGTAACACGAGTAATTAAGCAAAATCCGACGACAAAACTTCAACCACTCGAAATTC	237
Db	1311 RRR	1252
Oy	238 ACGCTTCCTACTAAAATTTTGAAAGGAATCGATCAATACCAACCCATTACCAAAAATACA	297
Db	1251 RRR	1192
Oy	298 TAATCAAAATATGCGAGATCGTACCTCGAAACTTTGCTTCAGTCGCGACAGAGAGAAA	357
Db	1191 RRR	1132
Oy	358 GGAAGATCGTGAGAAAAGGGGTTAGGGTTTAAGCTCAGACTTATTTGAGTAATGGG	417
Db	1131 RRR	1072
Oy	418 ACGGTGCACATTTTCCT 436	
Db	1071 RRRRATCGCAAGCTCCT 1053	

RESULT 15
US-09-949-016-14004
; sequence 14004, Application US/09949016

```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 14004
; LENGTH: 4529
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14004

```

```

Query Match      3.9%; Score 45.6; DB 4; Length 4529;
Best Local Similarity 46.3%; Pred. No. 0.11;
Matches 150; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

```

```

QY 670 TGTTCCTGCTTTTATTTTCAATGATTTTGTATTTTATCTTCAATGAAAT 729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1561 TATTTTATGATATATATATATATATATATATATATATATATATATAT 1620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 730 TTGGCTGATCTTGCAAACTCTTCGATCATATGCAATATAGTGAACACTG 789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1621 TTATATATTTTATGATATATATATATATATATATATATATATATAT 1680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 790 TAATTTGTGTGTAATTTGTAATTTAGATTCTATTCGCGTTTAAAGTAAT 849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1681 ATATATTTATATATATATATATATATATATATATATATATATATAT 1740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 850 TGTATCATGTTTAAACATTTGTAAGTATGATTAATAAAGATAATTAGTAG 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1741 TATATATATTTATATATATTTTATGATATATATATATATATATATAT 1800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 910 ATAACTGGAAGCAAAAATGAGATAGATACATTTGATTTTGTGATTTTGA 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1801 ATATTTTATATATATTTATATATATATATATATATATATATATATAT 1860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 970 GAGAGTGAAGTACGCGCATGAGA 993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1861 TAGAGATATATGATACATATGACATA 1884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: August 14, 2005, 02:10:22
 Job time : 238 secs

This Page Blank (uspto)

Db 1 CACAAACATACCTCAAAATCCAGACTCAGATCTACATTAATGCAACTTCATCATGAA 60
QY 61 AACATCAAAAAAGTCAAAAGTAACAAAATCAAGTATCGACACACAAAGCCAGTAA 120
Db 61 AACATCAAAAAAGTCAAAAGTAACAAAATCAAGTATCGACACACAAAGCCAGTAA 120
QY 121 GATGAAATTTAAGAAAGCTCAGCTAGAGTGGGCAAAATCTTCTAATCAAAACAG 180
Db 121 GATGAAATTTAAGAAAGCTCAGCTAGAGTGGGCAAAATCTTCTAATCAAAACAG 180
QY 121 GATGAAATTTAAGAAAGCTCAGCTAGAGTGGGCAAAATCTTCTAATCAAAACAG 180
Db 121 GATGAAATTTAAGAAAGCTCAGCTAGAGTGGGCAAAATCTTCTAATCAAAACAG 180
QY 181 TAAACACAGATTAATAGCAAAATCCGACAGAAAATCTCAACCCACTCCGAAATTCACG 240
Db 181 TAAACACAGATTAATAGCAAAATCCGACAGAAAATCTCAACCCACTCCGAAATTCACG 240
QY 241 TCTTCACTAAATTTTGGAAAGAAATGATCAATACCAACCCATTACCAAAATCATAA 300
Db 241 TCTTCACTAAATTTTGGAAAGAAATGATCAATACCAACCCATTACCAAAATCATAA 300
QY 301 TCAAAATGCGAGAAATCTGTAAGTCAAGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 TCAAAATGCGAGAAATCTGTAAGTCAAGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 AGATGCGAGAAAGGGGTTTAAAGTCAAGTGGGCAAGTGGGCAAGTGGGCAAGTGGG 420
Db 361 AGATGCGAGAAAGGGGTTTAAAGTCAAGTGGGCAAGTGGGCAAGTGGGCAAGTGGG 420
QY 421 GTGTCACTTTTCCGTTTGGAAATGAACTTGGGCTCAGCTTATGGAGTAAATGATAT 480
Db 421 GTGTCACTTTTCCGTTTGGAAATGAACTTGGGCTCAGCTTATGGAGTAAATGATAT 480
QY 481 TGATGGGCTTTCTAGTAAATACAAATATAGTTATGGGCTTAAATTAAGCCCATGT 540
Db 481 TGATGGGCTTTCTAGTAAATACAAATATAGTTATGGGCTTAAATTAAGCCCATGT 540
QY 541 TGGAAATTTTGAACATGTCTTGGCTCTAGTCTTAAACATGACAGCAAGTGGTGG 600
Db 541 TGGAAATTTTGAACATGTCTTGGCTCTAGTCTTAAACATGACAGCAAGTGGTGG 600
QY 601 AGACAAAGTGGAGCATATACAAATGATCAAAACAGCTAGTGGGCTGGCTGGCTCAT 660
Db 601 AGACAAAGTGGAGCATATACAAATGATCAAAACAGCTAGTGGGCTGGCTGGCTCAT 660
QY 661 GTGTCACTTTTCTAGTAAATACAAATATAGTTATGGGCTTAAATTAAGCCCATGT 540
Db 661 GTGTCACTTTTCTAGTAAATACAAATATAGTTATGGGCTTAAATTAAGCCCATGT 540
QY 721 ATACAAATTTTGGGCTGTATCTTGCACAACTCTTGCATCATATCGCAATATATGTAACA 780
Db 721 ATACAAATTTTGGGCTGTATCTTGCACAACTCTTGCATCATATCGCAATATATGTAACA 780
QY 781 CTGGTATCTAATTTGTGTGTTAAATGTTAAATTTAGATTTCTCCGGTTTAAAG 840
Db 781 CTGGTATCTAATTTGTGTGTTAAATGTTAAATTTAGATTTCTCCGGTTTAAAG 840
QY 841 TGAATTAATATGATATGATGTTAAACATTTGTAAGTATGATTAATTAATGATTAATTT 900
Db 841 TGAATTAATATGATATGATGTTAAACATTTGTAAGTATGATTAATTAATGATTAATTT 900
QY 901 AGTGAATGATTAAGCTGAAGCAAAATGAGATAGATCAATTTGTTGTTGTTTGG 960
Db 901 AGTGAATGATTAAGCTGAAGCAAAATGAGATAGATCAATTTGTTGTTGTTTGG 960
QY 961 ACATATGCGAGAGTGAAGTACGCGCATGAAAGATCAAGAGACATTTGCTCAGAGCTACA 1020
Db 961 ACATATGCGAGAGTGAAGTACGCGCATGAAAGATCAAGAGACATTTGCTCAGAGCTACA 1020
QY 1021 AGTGAATGATTAAGCTGAAGTACGCGCATGAAAGATCAAGAGACATTTGCTCAGAGCTACA 1080
Db 1021 AGTGAATGATTAAGCTGAAGTACGCGCATGAAAGATCAAGAGACATTTGCTCAGAGCTACA 1080
QY 1081 ACCAAGAGCTCACTGACATATATTAATCTCTCTAGTCCCGTTCTCTCATCATCATCT 1140
Db 1081 ACCAAGAGCTCACTGACATATATTAATCTCTCTAGTCCCGTTCTCTCATCATCATCT 1140

QY 1141 CTCACACAAACAAAAG 1158
Db 1141 CTCACACAAACAAAAG 1158
RESULT 2
US-09-998-059-13
; Sequence 13, Application US/09998059
; Publication No. US20030005485A1
; GENERAL INFORMATION:
; APPLICANT: Ohirorge, John B.
; APPLICANT: Benning, Christoph
; APPLICANT: Gao, Hongbo
; APPLICANT: Girke, Thomas
; APPLICANT: White, Joseph A.
; TITLE OF INVENTION: Plant Seed Specific Promoters
; FILE REFERENCE: MSU-06689
; CURRENT APPLICATION NUMBER: US/09/998,059
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/250,401
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1164
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-998-059-13
Query Match 100.0%; Score 1158; DB 10; Length 1164;
Best Local Similarity 100.0%; Pred. No. 4e-255;
Matches 1158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAACATACCTCAAAATCCAGACTCAGATCTACATTAATGCAACTTCATCATGAA 60
Db 1 CACAAACATACCTCAAAATCCAGACTCAGATCTACATTAATGCAACTTCATCATGAA 60
QY 61 AACATCAAAAAAGTCAAAAGTAACAAAATCAAGTATCGACACACAAAGCCAGTAA 120
Db 61 AACATCAAAAAAGTCAAAAGTAACAAAATCAAGTATCGACACACAAAGCCAGTAA 120
QY 121 GATGAAATTTAAGAAAGCTCAGCTAGAGTGGGCAAAATCTTCTAATCAAAACAG 180
Db 121 GATGAAATTTAAGAAAGCTCAGCTAGAGTGGGCAAAATCTTCTAATCAAAACAG 180
QY 181 TAAACACAGATTAATAGCAAAATCCGACAGAAAATCTCAACCCACTCCGAAATTCACG 240
Db 181 TAAACACAGATTAATAGCAAAATCCGACAGAAAATCTCAACCCACTCCGAAATTCACG 240
QY 241 TCTTCACTAAATTTTGGAAAGAAATGATCAATACCAACCCATTACCAAAATCATAA 300
Db 241 TCTTCACTAAATTTTGGAAAGAAATGATCAATACCAACCCATTACCAAAATCATAA 300
QY 301 TCAAAATGCGAGAAATCTGTAAGTCAAGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 TCAAAATGCGAGAAATCTGTAAGTCAAGTGGGCAAGAGAGAGAGAGAGAGAGAGAG 360
QY 361 AGATGCGAGAAAGGGGTTTAAAGTCAAGTGGGCAAGTGGGCAAGTGGGCAAGTGGG 420
Db 361 AGATGCGAGAAAGGGGTTTAAAGTCAAGTGGGCAAGTGGGCAAGTGGGCAAGTGGG 420
QY 421 GTGTCACTTTTCCGTTTGGAAATGAACTTGGGCTCAGCTTATGGAGTAAATGATAT 480
Db 421 GTGTCACTTTTCCGTTTGGAAATGAACTTGGGCTCAGCTTATGGAGTAAATGATAT 480
QY 481 TGATGGGCTTTCTAGTAAATACAAATATAGTTATGGGCTTAAATTAAGCCCATGT 540
Db 481 TGATGGGCTTTCTAGTAAATACAAATATAGTTATGGGCTTAAATTAAGCCCATGT 540
QY 541 TGGAAATTTTGAACATGTCTTGGCTCTAGTCTTAAACATGACAGCAAGTGGTGG 600
Db 541 TGGAAATTTTGAACATGTCTTGGCTCTAGTCTTAAACATGACAGCAAGTGGTGG 600

QY	601	GGACAAAGCGGAGATATACATGGATCAAAACAGCCGTAAGTGCAGCGCTGCATCAT	660
Db	601	AGACAAAGTCGAGCATATACATGGATCAAAACAGCCGTAAGTGCAGCGCTGCATCAT	660
QY	661	GTGTCACCTGTTCCTCGTTTTTTTTTAATTTTTCATAAGTCTTTTGTGTTATCTTCA	720
Db	661	GTGTCACCTGTTCCTCGTTTTTTTTTAATTTTTCATAAGTCTTTTGTGTTATCTTCA	720
QY	721	ATACAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCCCAATATACGTGAACA	780
Db	721	ATACAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCCCAATATACGTGAACA	780
QY	781	CTGGGATCTAATTTGTTGGTAAATGTTAAATTTGATCTAATCTCCGGTTTTAAAG	840
Db	781	CTGGGATCTAATTTGTTGGTAAATGTTAAATTTGATCTAATCTCCGGTTTTAAAG	840
QY	841	TGAATTTATGTATCATGTTAAACATGTTAGTAAGATGATATAATAATGATAAATTT	900
Db	841	TGAATTTATGTATCATGTTAAACATGTTAGTAAGATGATATAATAATGATAAATTT	900
QY	901	AGTTGATGGATTAACGTGAAGCAAAAATAGATAGATACATTTGATTTTGTGTAATTTTG	960
Db	901	AGTTGATGGATTAACGTGAAGCAAAAATAGATAGATACATTTGATTTTGTGTAATTTTG	960
QY	961	ACATATGGGAGATGAGCTACGGCGCATGGAAGATCAAGAGCACTTCTGAGGTCAAG	1020
Db	961	ACATATGGGAGATGAGCTACGGCGCATGGAAGATCAAGAGCACTTCTGAGGTCAAG	1020
QY	1021	AGTACGCGTATAAAAAGCTTAGACTGAAGTCCCACTGCAAACTTAATCTTAAGTGCCTCAA	1080
Db	1021	AGTACGCGTATAAAAAGCTTAGACTGAAGTCCCACTGCAAACTTAATCTTAAGTGCCTCAA	1080
QY	1081	ACCACGAGCTCACTTGACATATATATAAATCTCTTAAGTCCGTTCTCTTCATCCATCT	1140
Db	1081	ACCACGAGCTCACTTGACATATATATAAATCTCTTAAGTCCGTTCTCTTCATCCATCT	1140
QY	1141	CTCACAAACAAACAAAAG 1158	
Db	1141	CTCACAAACAAACAAAAG 1158	

```

RESULT 3
US-09-938-842A-4937
; Sequence 4937, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4937
; LENGTH: 1223
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4937

Query Match 100.0%; Score 1158; DB 9; Length 1223;
Best Local Similarity 100.0%; Pred. No. 4,1e-255;
Matches 1158; Conservative 0; Mismatches 0; Gaps 0;

```

OY	1	TAGCAACATACCTCAAAATCCACAATCCACATCTACCTCAATTAAGCAATTCATCAAGAA	60
Db	63	CACAAACATACCTCAAAATCCACAATCCACATCTACCTCAATTAAGCAATTCATCAAGAA	122
OY	61	AACATCAAAAAACGTCAAAGTAAACAAATCAAGTCAGATTACAGACAACAAAGCAGTAAA	120
Db	123	AACATCAAAAAACGTCAAAGTAAACAAATCAAGTCAGATTACAGACAACAAAGCAGTAAA	182
OY	121	GATGAAAATTTAAGCAACGCTCATGCTTAAGCTGCGCAAAATACTCTTAATCAAAAACG	180
Db	183	GATGAAAATTTAAGCAACGCTCATGCTTAAGCTGCGCAAAATACTCTTAATCAAAAACG	242
OY	181	TAACAAGAGAAATTACCAAAATCCGAGCGAAAACTCACCACCCAGCTCGAAATTACG	240
Db	243	TAACAAGAGAAATTACCAAAATCCGAGCGAAAACTCACCACCCAGCTCGAAATTACG	302
OY	241	TCTTCACATAAATTTTGAAAAGAAAGATGATCAATACCAACCCATTACACAAAATACATAA	300
Db	303	TCTTCACATAAATTTTGAAAAGAAATGATCAATACCAACCCATTACACAAAATACATAA	362
OY	301	TCAAAATGCGAGAAATCGTACCTGGAATCTTTGCTTCAAGTCGAGAGAGGAAAAAGAA	360
Db	363	TCAAAATGCGAGAAATCGTACCTGGAATCTTTGCTTCAAGTCGAGAGAGGAAAAAGAA	422
OY	361	AGATCGTGGAGAAAGGGGTTTAAGGTTTAAGCTCAGACTTCTAATGAGTAAAGGACG	420
Db	423	AGATCGTGGAGAAAGGGGTTTAAGGTTTAAGCTCAGACTTCTAATGAGTAAAGGACG	482
OY	421	GTGTCACATTTCCGTTTGGAAAATGAACCTTGGGGCTCAGCTTAATGAGGCTTAATGATTT	480
Db	483	GTGTCACATTTCCGTTTGGAAAATGAACCTTGGGGCTCAGCTTAATGAGGCTTAATGATTT	542
OY	481	TGATGGGCTTTCTAGTAATACAAATATAGTATTGGGCTTAAGTTTAATTAAGCCATGT	540
Db	543	TGATGGGCTTTCTAGTAATACAAATATAGTATTGGGCTTAAGTTTAATTAAGCCATGT	602
OY	541	TGGAATAATTTTGACACATGCTTGGGCTACTAGTCTAAACATGCAACCGAACATTTGTGG	600
Db	603	TGGAATAATTTTGACACATGCTTGGGCTACTAGTCTAAACATGCAACCGAACATTTGTGG	662
OY	601	AGACAAGTCGAGCATATACAAATGAGTCAAAACAGCTAGTGGCGCGCTCGCTCAT	660
Db	663	AGACAAGTCGAGCATATACAAATGAGTCAAAACAGCTAGTGGCGCGCTCGCTCAT	722
OY	661	GTGTCACCTGTGTTCCCTGTTTTTTTTTAAATTTTCAATAGTTCTTTTGTTTTATCTTCA	720
Db	723	GTGTCACCTGTGTTCCCTGTTTTTTTTTAAATTTTCAATAGTTCTTTTGTTTTATCTTCA	782
OY	721	ATACAAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATACGTGAACA	780
Db	783	ATACAAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATACGTGAACA	842
OY	781	CTGCGTATCTAATTTGTTGTGTTAAATGTTAAATTTAGATTCTAATCTCGGGTTTAAAG	840
Db	843	CTGCGTATCTAATTTGTTGTGTTAAATGTTAAATTTAGATTCTAATCTCGGGTTTAAAG	902
OY	841	TGAATTTATGATATCATGTTTAAACCACTTGAATGAAGATATATAATGATATAATTT	900
Db	903	TGAATTTATGATATCATGTTTAAACCACTTGAATGAAGATATATAATGATATAATTT	962
OY	901	AGTTGATGATATACGTGAAGCAAAAAATGAGATAGATACATTTGATTTTGTCTGATTTTG	960
Db	963	AGTTGATGATATACGTGAAGCAAAAAATGAGATAGATACATTTGATTTTGTCTGATTTTG	1022
OY	961	ACATATCGGAGAGTGAAGTACCGGCGATGGAAGATCAAGACACTTGTCTGAGCTCACAG	1020
Db	1023	ACATATCGGAGAGTGAAGTACCGGCGATGGAAGATCAAGACACTTGTCTGAGCTCACAG	1082
OY	1021	AGTAGACGTGTAAGAAAGCTTAGACATGAAGTCCCATGCAAACTTAATCTACGTGGCTCAA	1080
Db	1083	AGTAGACGTGTAAGAAAGCTTAGACATGAAGTCCCATGCAAACTTAATCTACGTGGCTCAA	1142
OY	1081	ACCAAGAGCTCATTTGACAAATATATATAATCTTCTTAAGTCCGGTTCTTTCAATCATCTT	1140

Db	663	AGACAAGTCGAGCATATACAAATGSAATCAACAGCGCTAATGTGTGCGCGCTCGTCAT	722
Qy	661	GTGTACACCTTGTTCCCGATTTTTTTTTTAATTTTCAATAGTCTTTTGTTTATCTTCA	720
Db	723	GTGTACACCTGTGTTCCCGATTTTTTTTTTAATTTTCAATAGTCTTTTGTTTATCTTCA	782
Qy	721	ATACAAATTTTTGGCTGTATCTTGCAAACTCTTGATCATATCGCAATATACGTGAACA	780
Db	783	ATACAAATTTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATACGTGAACA	842
Qy	781	CTGTGATCTTAATTTGTTGTGTTAATGTTAAATTTAGATTCTATTCCTCGGTTTAAAG	840
Db	843	CTGTGATCTTAATTTGTTGTGTTAATGTTAAATTTAGATTCTATTCCTCGGTTTAAAG	902
Qy	841	TGAATTTATGATCATGTTTAAACATGTGAAGTAGATGATTAATAAATGATAAATTT	900
Db	903	TGAATTTATGATCATGTTTAAACATGTGAAGTAGATGATTAATAAATGATAAATTT	962
Qy	901	AGTTGATGSAATAACGTGAGCAAAAATGAGATAGATACATTTGATTTTGTCTGATTTTG	960
Db	963	AGTTGATGSAATAACGTGAGCAAAAATGAGATAGATACATTTGATTTTGTCTGATTTTG	1022
Qy	961	ACATATGCGAGATGAGCTAACGGGCATGAAAGATCAAGAGACACTTGTCTGAGCTCACAG	1020
Db	1023	ACATATGCGAGATGAGCTAACGGGCATGAAAGATCAAGAGACACTTGTCTGAGCTCACAG	1082
Qy	1021	AGTACGCTGTAAAGGCTTAGACTGAAGTCCCATGCAAACTTAATCTACGTGTGCTCAA	1080
Db	1083	AGTACGCTGTAAAGGCTTAGACTGAAGTCCCATGCAAACTTAATCTACGTGTGCTCAA	1142
Qy	1081	ACCAAGAGCTCACTTGCAATATATATAAATCTTCTAATGCCGTTCTTTCAATCAATCT	1140

Db	1143	ACCACGAGCTC	ACTTGACAAATAT	TAACCTCTCTAAGTCCCGTCTCTTCATCATCT	1202
Qy	1141	CTCACAACAAAC	AAAAAG		1158
Db	1203	CTCACAACAAAC	AAAAAG		1220

RESULT 4

```

Sequence 4937: Application US/09938842A
Publication No. US20040009476A9
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: S13PI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 4937
LENGTH: 1223
TYPE: DNA
ORGANISM: Arabidopsis thaliana
IS-09-938-842A-4937

```

Query Match	100.0%	Score 1158	DB 11	Length 1223
Best Local Similarity	100.0%	Pred. No. 4.1e-255		
Matches 1158	Conservative 0	Mismatches 0	Indels 0	Gaps 0

Qy	1	CACAAACATACACTCAAAATCCAGACTCAATCTAAGCTAAATTAAGCAACTTCATCAAGAA	60
Db	63	CACAAACATACACTCAAAATCCAGACTCAATCTAAGCTAAATTAAGCAACTTCATCAAGAA	122
Qy	61	AACATCAAAAACAGCTCAAAAGTAACAAAATCAAAGTCAGATTCAGACAACAAAGCCAGTAAA	120
Db	123	AACATCAAAAACAGCTCAAAAGTAACAAAATCAAAGTCAGATTCAGACAACAAAGCCAGTAAA	182
Qy	121	GATAGAAAATTTAAGSAGCGCTCAGCTAAGCTGGCGAAAATACTTCTTAATCAAAAAG	180
Db	183	GATAGAAAATTTAAGSAGCGCTCAGCTAAGCTGGCGAAAATACTTCTTAATCAAAAAG	242
Qy	181	TAAACAAGAGTAATTAAGCAAAATCCGAGCAGAAAAGCTCACCCACCTCGGAATTCAG	240
Db	243	TAAACAAGAGTAATTAAGCAAAATCCGAGCAGAAAAGCTCACCCACCTCGGAATTCAG	302
Qy	241	TCTTCACTAAAAATTTTCGAAAGAAATGCATCAATACAAACCATTAACAAAAATACATA	300
Db	303	TCTTCACTAAAAATTTTCGAAAGAAATGCATCAATACAAACCATTAACAAAAATACATA	362
Qy	301	TCAAAATGGCGGAATCGTAACCTGGAAATCTTGCTCAAGTCGGAGAGAGGAAAAAGGA	360
Db	363	TCAAAATGGCGGAATCGTAACCTGGAAATCTTGCTCAAGTCGGAGAGAGGAAAAAGGA	422
Qy	361	AGATCGTGGAGAAAGGGGTTTAGGGTTTAAGCTCAGACTTCTAATGGAATGCGGACG	420
Db	423	AGATCGTGGAGAAAGGGGTTTAGGGTTTAAGCTCAGACTTCTAATGGAATGCGGACG	482
Qy	421	GTGTCACATTTTCCGTTTGGAAATGAACCTTGGGCTCAGCTTAAGGGCTATTAGATATT	480
Db	483	GTGTCACATTTTCCGTTTGGAAATGAACCTTGGGCTCAGCTTAAGGGCTATTAGATATT	542
Qy	481	TGATGGGCTTTCTAGTAAATACAAATTAAGTTATTGGGCTTAGTTAAATAGCCCATGT	540

Db	543	TGATGGGCTTTCCTGTAATAATACAATATAAAGTTATTTGGGCTAGTTTAAATTAAGCCCATGT	602
Qy	541	TGGAATATTTTGAACAATGCTTTGGCTATAGTGCCTAAACATGGAACCGAACAATTATGCG	600
Db	603	TGGAAATATTTTGAACAATGCTTTGGCTATAGTGCCTAAACATGGAACCGAACAATTATGCG	662
Qy	601	AGACAAGTCGAGCATATACATGATGATCAACACGCGCTAGTGTGCGCGCTTGCTGCAT	660
Db	663	AGACAAGTCGAGCATATACATGATGATCAACACGCGCTAGTGTGCGCGCTTGCTGCAT	722
Qy	661	GTGTCACCTTGTTCCCTCGTTTTTTTTTAATTTTTCATAAAGTCTTTTGTTTATCTTCA	720
Db	723	GTGTCACCTTGTTCCCTCGTTTTTTTTTAATTTTTCATAAAGTCTTTTGTTTATCTTCA	782
Qy	721	ATACAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATAGCTGAACA	780
Db	783	ATACAAATTTTGGCTGTATCTTGCAAACTCTTGATCATATGCGCAATATAGCTGAACA	842
Qy	781	CTGGTGATCTAATTTGTGTGTTTAATGTTAAATTAAGTTTATCTCTCGGTTTTAAAG	840
Db	843	CTGGTGATCTAATTTGTGTGTTTAATGTTAAATTAAGTTTATCTCTCGGTTTTAAAG	902
Qy	841	TGAATTTATATGATCATAGTTTAAACATGTAAGTAAAGATGATATATATTAAGTAAATTT	900
Db	903	TGAATTTATATGATCATAGTTTAAACATGTAAGTAAAGATGATATATATTAAGTAAATTT	962
Qy	901	AGTTGATGATTAACGTGAAGCAAAAAATGAGATGATATACATTTGATTTTGTGCTATTTTG	960
Db	963	AGTTGATGATTAACGTGAAGCAAAAAATGAGATGATATACATTTGATTTTGTGCTATTTTG	1022
Qy	961	ACATATGCGGAGATGAGCTACGCGCATGAAAGATCAAGAGACATTTGCTGAGCTCACAG	1020
Db	1023	ACATATGCGGAGATGAGCTACGCGCATGAAAGATCAAGAGACATTTGCTGAGCTCACAG	1082
Qy	1021	AGTACGCTGTAATAAAGCTTAGATGCTGAAGTCCCGCATGCAAAACCTATCTTACGTGCTCAA	1080
Db	1083	AGTACGCTGTAATAAAGCTTAGATGCTGAAGTCCCGCATGCAAAACCTATCTTACGTGCTCAA	1142
Qy	1081	ACGACGAGCTCACTTGACAATATATTAACCTCTCTAAGTCCGCTTCTTTGATCATCACTT	1140
Db	1143	ACGACGAGCTCACTTGACAATATATTAACCTCTCTAAGTCCGCTTCTTTGATCATCACTT	1202
Qy	1141	CTGCACAACAAACAAAAAG 1158	
Db	1203	CTGCACAACAAACAAAAAG 1220	

RESULT

```

US-10-892-513-4
Sequence 4, Application US/10892513
Publication No. US20050039229A1
GENERAL INFORMATION:
APPLICANT: UNICROD LTD
TITLE OF INVENTION: Double recoverable block of function
FILE REFERENCE: Kuvshinov double construct
CURRENT APPLICATION NUMBER: US/10/892,513
PRIOR FILING DATE: 2004-07-15
PRIORITY APPLICATION NUMBER: 09/617,543
PRIOR FILING DATE: 2000-07-14
NUMBER OF SEQ. ID NOS: 16
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 1183
TYPE: DNA
ORGANISM: artificial
FEATURE:
OTHER INFORMATION: Brassica napus
NAME/KEY: misc_feature
OTHER INFORMATION: CRU promoter with artificial polyadenylation site
US-10-892-513-4

```

Query Match 17.2%; Score 198.8; DB 21; Length 1183;

Best Local Similarity 66.0%; Pred. No. 3.4e-35;
Matches 409; Conservative 0; Mismatches 177; Indels 34; Gaps 7;
QY 568 ACTAGTCTTAAACATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 627
DB 567 ACTGCTCTTGAATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 626
QY 628 CAACAAGCTTGTGAGACAGTGGAGATATACAAATGAT 687
DB 627 CAACAAGCTTGTGAGACAGTGGAGATATACAAATGAT 682
QY 688 TAATTTTCTAAGTCTTGTGAGACAGTGGAGATATACAAATGAT 747
DB 683 TAATTTTCTAAGTCTTGTGAGACAGTGGAGATATACAAATGAT 742
QY 748 ACTTGTGATATATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 801
DB 743 ACTTGTGATATATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 802
QY 802 TTAATTTGT-TAAATTTAGATTTCTGCGTTTAAAGTGAATATATATGATGAT 860
DB 803 TTAATTTGT-TAAATTTAGATTTCTGCGTTTAAAGTGAATATATATGATGAT 862
QY 861 TAAACATTTAGTAAAGTATATATATATATATATATATATATATATATATAT 920
DB 863 TGAATTTAGTAAAGTATATATATATATATATATATATATATATATATATAT 922
QY 921 -----CAAAAATGATAGATATATATATATATATATATATATATATATAT 959
DB 923 CTGATTAACATGAAACATTTTCTTAAATAAGATATATATATATATATATATAT 982
QY 960 GACATATGCGAGAGTGAAGTACGCGCATGAGATCAAGACATTTGCTGAGCTACA 1019
DB 983 GAGATATGCGAGAGTGAAGTACGCGCATGAGATCAAGACATTTGCTGAGCTACA 1042
QY 1020 GAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1079
DB 1043 GAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1102
QY 1080 AACACAGAGCTACTGCAATATATATATATATATATATATATATATATATAT 1137
DB 1103 AACCTTCACTGCTACTGCAATATATATATATATATATATATATATATATAT 1162
QY 1138 TCTCTCACAACAAACAAA 1157
DB 1163 TCTATCACAACAAACAAA 1182
RESULT 6
US-10-892-513-10
; Sequence 10, Application US/10892513
; Publication No. US20050039229A1
; GENERAL INFORMATION:
; APPLICANT: Unilever Ltd
; TITLE OF INVENTION: Double recoverable block of function
; FILE REFERENCE: Kuvshinov double construct
; CURRENT APPLICATION NUMBER: US/10/892,513
; PRIOR FILING DATE: 2004-07-15
; PRIOR APPLICATION NUMBER: 09/617,543
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 1293
; TYPE: DNA
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Braesica napus
; NAME/KEY: misc feature
; LOCATION: (1)-(1239)
; OTHER INFORMATION: Changed CRU promoter with three test operators in the vicinity of
; OTHER INFORMATION: TATA box

FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1226)-(1244)
; OTHER INFORMATION: Tet operator
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1246)-(1252)
; OTHER INFORMATION: TATA-box
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1254)-(1272)
; OTHER INFORMATION: Tet operator
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1275)-(1293)
; OTHER INFORMATION: Tet operator
US-10-892-513-10
Query Match 17.2%; Score 198.8; DB 21; Length 1293;
Best Local Similarity 66.0%; Pred. No. 3.6e-35;
Matches 409; Conservative 0; Mismatches 177; Indels 34; Gaps 7;
QY 568 ACTAGTCTTAAACATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 627
DB 567 ACTGCTCTTGAATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 626
QY 628 CAACAAGCTTGTGAGACAGTGGAGATATACAAATGAT 687
DB 627 CAACAAGCTTGTGAGACAGTGGAGATATACAAATGAT 682
QY 688 TAATTTTCTAAGTCTTGTGAGACAGTGGAGATATACAAATGAT 747
DB 683 TAATTTTCTAAGTCTTGTGAGACAGTGGAGATATACAAATGAT 742
QY 748 ACTTGTGATATATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 801
DB 743 ACTTGTGATATATGCAACCGAAGCTTGTGAGACAGTGGAGATATACAAATGAT 802
QY 802 TTAATTTGT-TAAATTTAGATTTCTGCGTTTAAAGTGAATATATATATATATAT 860
DB 803 TTAATTTGT-TAAATTTAGATTTCTGCGTTTAAAGTGAATATATATATATATAT 862
QY 861 TAAACATTTAGTAAAGTATATATATATATATATATATATATATATATATATAT 920
DB 863 TGAATTTAGTAAAGTATATATATATATATATATATATATATATATATATAT 922
QY 921 -----CAAAAATGATAGATATATATATATATATATATATATATATATAT 959
DB 923 CTGATTAACATGAAACATTTTCTTAAATAAGATATATATATATATATATATAT 982
QY 960 GACATATGCGAGAGTGAAGTACGCGCATGAGATCAAGACATTTGCTGAGCTACA 1019
DB 983 GAGATATGCGAGAGTGAAGTACGCGCATGAGATCAAGACATTTGCTGAGCTACA 1042
QY 1020 GAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1079
DB 1043 GAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1102
QY 1080 AACACAGAGCTACTGCAATATATATATATATATATATATATATATATATAT 1137
DB 1103 AACCTTCACTGCTACTGCAATATATATATATATATATATATATATATATAT 1162
QY 1138 TCTCTCACAACAAACAAA 1157
DB 1163 TCTATCACAACAAACAAA 1182
RESULT 7
US-10-311-455-2139/c
; Sequence 2139, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander

APPLICANT: PIERPNEROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
TITLE OF INVENTION: Cytosine methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 2139
LENGTH: 10326
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-2139

Query Match
Best Local Similarity 5.6%; Score 65.4; DB 15; Length 10326;
Matches 177; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

QY 2 ACAAACATGACTCAAAATCCAGACTCATCTACTCAATTATGCAACTTCATCATGAA 61
DB 2764 ACCAATCACTCAAAACCAACCAATTAATTAACCTTAACCAAAATTTCAACAT 2705
QY 62 ACATCAAAAACGTCAGAAATCAAAATCAAGTCAGATTCAGACACAAAGCAGTAAAG 121
DB 2704 TCACCTCAAAACACAAATTTTCTACCTAATAAATAACAAACAAACAAATTA 2645
QY 122 ATAGAAATTTTAAAGAGCTCATGCTAAGCTGCGCAAAATCTCTAATCAAAACAGT 181
DB 2644 AAAAAATTTTAAACACCTCCAAAAAATATTAACAAACACAAATCTCAACATATA 2585
QY 182 AACAAAGATTAATAGCAAAATCCGAGCAAAACCTCCACCCACCTCGAAATTCAGT 241
DB 2584 TACGATTCGCAATTTCTAAATTAATAAAGCAATTAACAAACAAATTAATACTA 2525
QY 242 CTTCCTAAATTTTTCGAAAGATCGATCATACCAACCTTACACAAATTAATAT 301
DB 2524 TATCACTCCCTACTTAAAAAATAATCAATAAATAATTAATTCACAAATTAATAT 2465
QY 302 CAAAATGGGAGAAATCGTACTGGAACCTTCTCAAGTCGAGAGAGAGAAAGGA 361
DB 2464 TAAATAATAACAAATTTTAAACAAATATCTCTTACTACTCAATATATATAAAAAA 2405
QY 362 GAT 364
DB 2404 TAT 2402

RESULT 8
US-10-021-323-16890/c
Sequence 16890, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plantes
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 16890

LENGTH: 547
TYPE: DNA
ORGANISM: Goseyplum hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3829-025-Q6-N6-C2
US-10-021-323-16890

Query Match
Best Local Similarity 5.5%; Score 64.2; DB 19; Length 547;
Matches 138; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 680 TTTTCTTTTAAATTTTTCATAGTCTTTTGTATATCTGCAATCAATTTTGGCGTA 739
DB 350 TTTTCTTTTAAATTTTTCATAGTCTTTTGTATATCTGCAATCAATTTTGGCGTA 739
QY 740 TCTTGAACCTCTTCATATATGATGATGATGATGATGATGATGATGATGATGATG 799
DB 290 TTTTAAATTTTCCCTTTTCCCTTTTCCCTTTTCCCTTTTCCCTTTTCCCTTTT 231
QY 800 TGTATATGTTAAATTTAGATCTATCTCTCGGTTTAAAGTGAATATATATGATG 859
DB 230 TTTTCTTTTAACTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 171
QY 860 TTAACATCTGTATAGTATGATGATGATGATGATGATGATGATGATGATGATG 919
DB 170 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 111
QY 920 GCAAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
DB 110 AAAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 90

RESULT 9
US-10-021-323-6681/c
Sequence 6681, Application US/10021323
Publication No. US20040123340A1
GENERAL INFORMATION:
APPLICANT: Deikman, Jill
APPLICANT: Feng, Paul C.C.
APPLICANT: Fincher, Karen L.
APPLICANT: Ziegler, Todd E.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plantes
FILE REFERENCE: 38-21(52274)B
CURRENT APPLICATION NUMBER: US/10/021,323
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: US 60/255, 619
PRIOR FILING DATE: 2000-12-14
NUMBER OF SEQ ID NOS: 17880
SEQ ID NO 6681
LENGTH: 433
TYPE: DNA
ORGANISM: Goseyplum hirsutum
FEATURE:
OTHER INFORMATION: Clone ID: LIB3828-009-Q1-N6-D11
US-10-021-323-6681

Query Match
Best Local Similarity 4.9%; Score 56.4; DB 19; Length 433;
Matches 141; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 657 TCATGTGACCTGTTCTCTGCTTTTAAATTTTCAATGATCTTTGTTATATC 716
DB 282 TTATGTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 223
QY 717 TTCAATACAAATTTTGGCTATCTTGCAACCTTCGATCATATGCGCAATATAGTG 766
DB 222 TTTTAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 163
QY 777 AACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 836
DB 162 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTT 103

; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 2424
 ; SEQ ID NO 1053
 ; LENGTH: 6713
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-311-455-1053

Query Match 4.7%; Score 54.2; DB 15; Length 6713;
 Best Local Similarity 49.1%; Pred. No. 0.11;
 Matches 143; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 16 AAATCCAGACTCACATCTTACTCAATTATGCACTTCATCATGAAAAACATCAAAAACAGT 75
 DB 1045 AAAATATCACTCCCTCACTCAATTAATACTTATATATAAAAAACAATAATTAACCGA 986
 QY 76 CAAAGTACAAAATCAAGTCAGATTCAGACACAAAAGCCAGTAAGTAGAAAATTAAAC 135
 DB 985 TAAACAAATTAATAATCAAAAAAACCCCTCATACATATCTAATAATATTAATTAC 926
 QY 136 GAACGCTCATGTAGCTGCGCAAAATACTTCTTAATCAAAACAGTAACAACGAGTAATT 195
 DB 925 AACCACTATTAATAAACAAATAATAAAATCTCAAAAAATAAAAGATTAATA 866
 QY 196 AGCAAAATCCGAGCAGAAAACTCTCACCCCTCCGAAATTCAGAGTCTTCACTAAATT 255
 DB 865 ATCTAACAAATTTTATTAATAATATATCTTAATAAAAAAATCAATTAATTAATAATA 806
 QY 256 TCGAAGAAGATCGATCAATACCAACCATTAACAAAAATACATATCAAAA 306
 DB 805 TCTACACTCCCATTTTATTACACACATATTCACAAATTAATTAATAATAATA 755

Search completed: August 14, 2005, 04:24:34
 Job time : 890 secs

This Page Blank (uspto)